

03560.003438.



PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:	)	
	:	Examiner: Not Yet Assigned
HIROTO YOSHII, ET AL.	)	
	:	Group Art Unit: Not Yet Assigned
Application No.: 10/805,292	)	
	:	
Filed: March 22, 2004	)	
	:	
For: DNA PROBE DESIGN DEVICE AND	)	
INFORMATION PROCESSING	:	
METHOD FOR DNA PROBE DESIGN	)	April 29, 2004

Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

SUBMISSION OF PRIORITY DOCUMENTS

Sir:

In support of Applicants' claim for priority under 35 U.S.C. § 119, enclosed  
are certified copies of the following foreign applications:

2004-064494, filed March 8, 2004; and

2003-099464, filed April 2, 2003.

Applicants' undersigned attorney may be reached in our New York office by telephone at (212) 218-2100. All correspondence should continue to be directed to our address given below.

Respectfully submitted,

Peter C. Theulius  
Attorney for Applicants

Registration No. 47,138

FITZPATRICK, CELLA, HARPER & SCINTO  
30 Rockefeller Plaza  
New York, New York 10112-3800  
Facsimile: (212) 218-2200

NY\_MAIN 424547v1

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日                      2 0 0 4 年    3 月    8 日  
Date of Application:

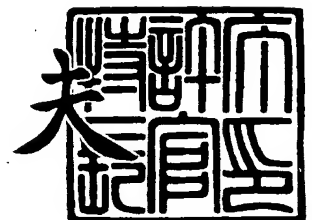
出 願 番 号                      特 願 2 0 0 4 - 0 6 4 4 9 4  
Application Number:  
[ST. 10/C]:                      [ J P 2 0 0 4 - 0 6 4 4 9 4 ]

出      願      人                      キヤノン株式会社  
Applicant(s):

2 0 0 4 年    4 月 1 9 日

特許庁長官  
Commissioner,  
Japan Patent Office

今 井 康 夫



出証番号    出証特 2 0 0 4 - 3 0 3 2 9 0 6

【書類名】 特許願  
【整理番号】 5516222-01  
【提出日】 平成16年 3月 8日  
【あて先】 特許庁長官殿  
【国際特許分類】 G06F 3/00  
【発明者】  
    【住所又は居所】 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内  
    【氏名】 吉井 裕人  
【発明者】  
    【住所又は居所】 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内  
    【氏名】 福井 寿文  
【特許出願人】  
    【識別番号】 000001007  
    【氏名又は名称】 キヤノン株式会社  
【代理人】  
    【識別番号】 100076428  
    【弁理士】  
    【氏名又は名称】 大塚 康徳  
    【電話番号】 03-5276-3241  
【選任した代理人】  
    【識別番号】 100112508  
    【弁理士】  
    【氏名又は名称】 高柳 司郎  
    【電話番号】 03-5276-3241  
【選任した代理人】  
    【識別番号】 100115071  
    【弁理士】  
    【氏名又は名称】 大塚 康弘  
    【電話番号】 03-5276-3241  
【選任した代理人】  
    【識別番号】 100116894  
    【弁理士】  
    【氏名又は名称】 木村 秀二  
    【電話番号】 03-5276-3241  
【先の出願に基づく優先権主張】  
    【出願番号】 特願2003- 99464  
    【出願日】 平成15年 4月 2日  
【手数料の表示】  
    【予納台帳番号】 003458  
    【納付金額】 21,000円  
【提出物件の目録】  
    【物件名】 特許請求の範囲 1  
    【物件名】 明細書 1  
    【物件名】 図面 1  
    【物件名】 要約書 1  
    【包括委任状番号】 0102485

**【書類名】 特許請求の範囲****【請求項 1】**

ターゲットの塩基配列データを含む第 1 の塩基配列データ群について、該ターゲットの塩基配列データから得られる複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第 1 カウント工程と、

前記第 1 の塩基配列データ群とは区別されるべき第 2 の塩基配列データ群について、前記複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第 2 カウント工程と、

前記第 1 及び第 2 カウント工程によって保持された頻度情報を、前記ターゲットの塩基配列を基準に対比可能に表示する表示工程と、

ユーザによる指示操作に応じて前記複数の部分塩基配列の少なくとも一つを抽出し、抽出された部分塩基配列に基づいてプローブ候補を決定する決定工程とを備えることを特徴とする DNA プローブ設計のための情報処理方法。

**【請求項 2】**

前記表示工程は前記第 1 及び第 2 カウント工程によって保持された頻度情報を、前記ターゲットの塩基配列を 1 つの軸とするグラフで表示し、

前記指示操作は、前記グラフ上の所望の位置を指示することによりなされることを特徴とする請求項 1 に記載の DNA プローブ設計のための情報処理方法。

**【請求項 3】**

前記第 1 の塩基配列データ群と前記第 2 の塩基配列データ群に含まれる全ての塩基配列データに共通して存在する共通配列を抽出する抽出工程をさらに備え、

前記表示工程は、前記グラフ上に、前記抽出工程で抽出された共通配列領域を示すことを特徴とする請求項 2 に記載の DNA プローブ設計のための情報処理方法。

**【請求項 4】**

前記決定工程において、前記共通配列領域ではユーザによる部分塩基配列の指示を行えないようにすることを特徴とする請求項 3 に記載の DNA プローブ設計のための情報処理方法。

**【請求項 5】**

ターゲットの塩基配列データを含む第 1 の塩基配列データ群について、該ターゲットの塩基配列データから得られる複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第 1 カウント工程と、

前記第 1 の塩基配列データ群とは区別されるべき第 2 の塩基配列データ群について、前記複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第 2 カウント工程と、

前記第 1 カウント工程で得られた頻度と前記第 2 カウント工程で得られた頻度に基づいて、前記ターゲットの塩基配列から部分塩基配列を抽出してプローブ候補を決定する決定工程とを備えることを特徴とする DNA プローブ設計のための情報処理方法。

**【請求項 6】**

前記決定工程は、前記第 1 カウント工程による頻度が第 1 所定値を超え、前記第 2 カウント工程による頻度が第 2 所定値を下回る部分塩基配列を前記ターゲットの塩基配列より抽出することを特徴とする請求項 5 に記載の DNA プローブ設計のための情報処理方法。

**【請求項 7】**

前記第 1 の塩基配列データ群と前記第 2 の塩基配列データ群に含まれる全ての塩基配列データに共通して存在する共通配列を抽出する抽出工程をさらに備え、

前記決定工程は、前記ターゲットの塩基配列の、前記抽出工程で抽出された共通配列からなる領域を除く領域より部分塩基配列を抽出することを特徴とする請求項 5 に記載の DNA プローブ設計のための情報処理方法。

**【請求項 8】**

前記決定工程は、抽出された部分塩基配列について融解温度を算出し、算出された融解

温度に基づいてプローブ候補を決定することを特徴とする請求項 1 乃至 7 のいずれかに記載のプローブ設計のための情報処理方法。

【請求項 9】

前記決定工程は、抽出された部分塩基配列について、その前後に所定数の塩基を追加、削除することにより複数の部分塩基配列を派生し、該抽出及び派生された部分塩基配列の中からプローブセットに用いるべきプローブ候補を決定することを特徴とする請求項 1 乃至 8 のいずれかに記載の DNA プローブ設計方法。

【請求項 10】

前記決定工程は、抽出された部分塩基配列について 2 次構造が形成される可能性を算出し、その算出結果に基づいてプローブセットに用いるべきプローブ候補を決定することを特徴とする請求項 1 乃至 9 のいずれかに記載のプローブ設計のための情報処理方法。

【請求項 11】

前記決定工程は、抽出された部分塩基配列の相互のマッチング度を計算し、類似する部分塩基配列を排除することによりプローブ候補を決定することを特徴とする請求項 1 乃至 10 のいずれかに記載の DNA プローブ設計のための情報処理方法。

【請求項 12】

前記決定工程は、前記ターゲットの塩基配列における、前記共通領域によって区切られた部分領域を単位として、繰り返し返されることを特徴とする請求項 7 に記載の DNA プローブ設計のための情報処理方法。

【請求項 13】

前記決定工程は、前記共通領域によって区切られた複数の部分領域の各々から 1 つ以下の部分塩基配列を抽出することを特徴とする請求項 12 に記載の DNA プローブ設計のための情報処理方法。

【請求項 14】

前記決定工程は、前記複数の部分領域の各々より、前記第 1 カウント工程による頻度が最大であり、かつ、前記第 2 カウント工程による頻度が所定値以下となっている部分塩基配列を抽出することを特徴とする請求項 13 に記載の DNA プローブ設計のための情報処理方法。

【請求項 15】

前記決定工程は、前記複数の部分領域の各々より、前記第 1 カウント工程による頻度が所定値以上であり、かつ、前記第 2 カウント工程による頻度が所定値以下の部分塩基配列を抽出することを特徴とする請求項 12 に記載の DNA プローブ設計のための情報処理方法。

【請求項 16】

前記複数の部分塩基配列は、前記ターゲットの塩基配列データより所定数の塩基を抽出してなる塩基配列を、その先頭位置を順次ずらしながら取得することにより得られたものであることを特徴とする請求項 1 乃至 15 のいずれかに記載の DNA プローブ設計のための情報処理方法。

【請求項 17】

前記第 1 グループの塩基配列データ群はターゲットとする生物種の複数の多型を含む塩基配列データであり、前記第 2 グループの塩基配列データは前記ターゲットとする生物種以外の生物種の複数の多型を含む塩基配列データであることを特徴とする請求項 1 乃至 16 のいずれかに記載の DNA プローブ設計のための情報処理方法。

【請求項 18】

ターゲットの塩基配列データを含む第 1 の塩基配列データ群について、該ターゲットの塩基配列データから得られる複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第 1 カウント手段と、

前記第 1 の塩基配列データ群とは区別されるべき第 2 の塩基配列データ群について、前記複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第 2 カウント手段と、

前記第 1 及び第 2 カウント手段によって保持された頻度情報を、前記ターゲットの塩基配列を基準に対比可能に表示する表示手段と、

ユーザによる指示操作に応じて前記複数の部分塩基配列の少なくとも一つを抽出し、抽出された部分塩基配列に基づいてプローブ候補を決定する決定手段とを備えることを特徴とする DNA プローブ設計装置。

【請求項 19】

前記第 1 の塩基配列データ群と前記第 2 の塩基配列データ群に含まれる全ての塩基配列データに共通して存在する共通配列を抽出する抽出手段をさらに備え、

前記表示手段は前記第 1 及び第 2 カウント手段によって保持された頻度情報を前記ターゲットの塩基配列を 1 つの軸とするグラフで表示するとともに、該グラフ上に前記抽出手段で抽出された共通配列の位置を示すことを特徴とする請求項 18 に記載の DNA プローブ設計装置。

【請求項 20】

ターゲットの塩基配列データを含む第 1 の塩基配列データ群について、該ターゲットの塩基配列データから得られる複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第 1 カウント手段と、

前記第 1 の塩基配列群とは区別されるべき第 2 の塩基配列データ群について、前記複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第 2 カウント手段と、

前記第 1 カウント手段で得られた頻度と前記第 2 カウント手段で得られた頻度に基づいて、前記ターゲットの塩基配列から部分塩基配列を抽出してプローブ候補を決定する決定手段とを備えることを特徴とする DNA プローブ設計装置。

【請求項 21】

前記第 1 の塩基配列データ群と前記第 2 の塩基配列データ群に含まれる全ての塩基配列データに共通して存在する共通配列を抽出する抽出手段をさらに備え、

前記決定手段は、前記ターゲットの塩基配列にける、前記抽出手段で抽出された共通配列からなる領域を除く領域より部分塩基配列を抽出することを特徴とする請求項 20 に記載の DNA プローブ設計装置。

【請求項 22】

請求項 1 乃至 17 のいずれかに記載のプローブ設計のための情報処理方法をコンピュータに実行させるための制御プログラム。

【請求項 23】

請求項 1 乃至 17 のいずれかに記載のプローブ設計のための情報処理方法をコンピュータに実行させるための制御プログラムを格納した記憶媒体。

【請求項 24】

請求項 1 乃至 17 のいずれかに記載のプローブ設計方法で設計された核酸プローブを含む DNA マイクロアレイ。

【請求項 25】

請求項 24 に記載の DNA マイクロアレイを用いた核酸検査方法。

**【書類名】 明細書****【発明の名称】** D N A プローブ設計装置及び D N A プローブ設計のための情報処理方法**【技術分野】****【 0 0 0 1 】**

本発明はいわゆる D N A マイクロアレイ等を用いた核酸配列解析システムに好適なオリゴヌクレオチドプローブの設計技術に関するものである。

**【背景技術】****【 0 0 0 2 】**

従来より、特許文献 1 または特許文献 2 にあるように、D N A マイクロアレイを用いて遺伝子の発現やシーケンスを決定するシステムが知られている。これらの特許文献に開示されたシステムは、c D N A をスポットして作成される D N A マイクロアレイとは異なり、試料とハイブリダイズさせるためのプローブをあらかじめ設計する必要がある。うまく適切なプローブを設計できれば、非常に高い確率で試料の中に存在する塩基配列の断片に対する情報が得られる。

**【 0 0 0 3 】**

このシステムにおいて、プローブとして用いられる塩基配列の長さは、長くても 1 0 0 ベース以上になることは珍しく、短いものでは数個の塩基配列からなるものまでである。すなわち、上記特許文献で開示されたシステムでは、c D N A に比べて非常に短い塩基配列のプローブを用いて特定の塩基配列を捕らえることになる。このため、D N A 中のプローブとして用いる部分の塩基配列の特異性が非常に高い必要がある。

**【 0 0 0 4 】**

上記のような特異性の高い部分を選び出すための従来の選択方法では、一般的な配列に対する特異性評価が行われる。例えば、人間のゲノム由来の D N A を対象とした D N A マイクロアレイを作成しようとした場合、人間の全ゲノム塩基配列に対する特異性を調べて、特異性の高い部分をプローブ塩基配列として選んでいた。

**【特許文献 1】** 特開平 1 0 - 2 7 2 0 0 0 号公報**【特許文献 2】** 特開平 1 1 - 1 8 7 9 0 0 号公報**【発明の開示】****【発明が解決しようとする課題】****【 0 0 0 5 】**

しかしながら、従来の選択方法では、非常に似ている塩基配列が検体の中に入っていて、かつ、それら類似した塩基配列があるグループに属する塩基配列と別のグループに属する塩基配列を含む場合、各塩基配列がそのグループに属するかを判定することは非常に困難であるという問題が存在していた。より具体的には、感染症などの判定を行う際に、同種の菌の別種類の株の D N A 塩基配列には同じようなハイブリ強度を示し、かつ、別の種の菌の D N A 塩基配列には異なるハイブリ強度を示すようなプローブを見つけ出すことが非常に困難であるという問題があった。

**【 0 0 0 6 】**

また、従来の方法では、非常に似ている塩基配列の中で、生物種に特異的な場所、かつ、多型には共通な場所を探そうとする場合、マルチプルアラインメントなどの方法を用いて、対象としている生物種の全ての多型の塩基配列を揃えて表示し、人間が目で見ても妥当な部分を選んでいた。この従来法では、人間のミスが混入したり、人によるバラツキが生じていた。

**【 0 0 0 7 】**

本発明は、上記課題に鑑みてなされたものであり、正確かつ、再現性のあるプローブ設計を可能とすることを目的とする。

**【課題を解決するための手段】****【 0 0 0 8 】**

上記の目的を達成するための本発明による D N A プローブ設計のための情報処理方法は以下の工程を備える。即ち、



ターゲットの塩基配列データを含む第1の塩基配列データ群について、該ターゲットの塩基配列データから得られる複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第1カウント工程と、

前記第1の塩基配列データ群とは区別されるべき第2の塩基配列データ群について、前記複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第2カウント工程と、

前記第1及び第2カウント工程によって保持された頻度情報を、前記ターゲットの塩基配列を基準に対比可能に表示する表示工程と、

ユーザによる指示操作に応じて前記複数の部分塩基配列の少なくとも一つを抽出し、抽出された部分塩基配列に基づいてプローブ候補を決定する決定工程とを備える。

#### 【0009】

また、上記の目的を達成するための本発明の態様によるDNAプローブ設計のための情報処理方法は、

ターゲットの塩基配列データを含む第1の塩基配列データ群について、該ターゲットの塩基配列データから得られる複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第1カウント工程と、

前記第1の塩基配列群とは区別されるべき第2の塩基配列データ群について、前記複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第2カウント工程と、

前記第1カウント工程で得られた頻度と前記第2カウント工程で得られた頻度に基づいて、前記ターゲットの塩基配列から部分塩基配列を抽出してプローブ候補を決定する決定工程とを備える。

#### 【発明の効果】

#### 【0010】

本発明によれば、正確かつ、再現性のあるプローブ設計が可能となる。

#### 【発明を実施するための最良の形態】

#### 【0011】

以下、添付の図面に基づいて本発明の好適な実施形態を説明する。

#### 【0012】

<第1実施形態>

[プローブ設計方法の説明]

図2は第1実施形態によるプローブ設計方法が適用される情報処理装置の構成を示すブロック図である。本実施形態のプローブ設計方法は、外部記憶装置201、中央処理装置(CPU)202、メモリ203、入出力装置204から構成される装置に実装される。すなわち、一般的なパーソナルコンピュータ、ワークステーション等にも実装可能である。

#### 【0013】

図2において、外部記憶装置201は、本実施形態のプローブ設計方法を実現するプログラムや、各種塩基配列データ及びパラメータ(DNA(オリゴヌクレオチド)プローブ長や融解温度等)を保持する。また、本実施形態によって選択されたプローブ配列そのものを保持する機能を持つ。中央処理装置(CPU)202はプローブ設計のプログラムを実行したり、すべての装置の制御を行ったりする。メモリ203は中央処理装置(CPU)202が使用するプログラム、及びサブルーチンやデータを一時的に記憶する。入出力装置204は、ディスプレイ、キーボード、ポインティングデバイス等を含み、ユーザとのインタラクションを行う。多くの場合、本実施形態のプローブ設計方法を実現するプログラム実行のトリガはこの入出力装置を介してユーザが出す。また、ユーザが結果を見たり、プログラムのパラメータ制御をこの入出力装置を介して行う。

#### 【0014】

図1は、第1実施形態によるプローブ設計方法の処理手順を説明するフローチャートである。101は自己塩基配列データであり、例えばターゲット塩基配列107がある特定の菌の特定株の塩基配列であった場合、ターゲット塩基配列107と同種の菌の種々の株

が有する塩基配列のデータからなる。102は自己頻度テーブル作成ステップであり、自己塩基配列データ101に含まれる部分塩基配列の頻度をカウントして自己頻度テーブルを作成する。103はターゲット塩基配列と競合する競合塩基配列データであり、自己塩基配列データ101の菌種とは異なる種々の菌種の種々の株が有する塩基配列のデータからなる。104は競合頻度テーブル作成ステップであり、競合塩基配列データ103に含まれる部分塩基配列の頻度をカウントして競合頻度テーブルを作成するステップである。

#### 【0015】

108はプローブ評価ステップであり、自己頻度テーブル105と競合頻度テーブル106を用いて、ターゲット塩基配列107からプローブ候補を導き出す。109はプローブセット選択ステップであり、プローブ評価ステップ108の結果得られたプローブ候補から、より適切なプローブ候補のセットを選択する。なお、プローブセット選択ステップ110におけるプローブ候補の選別方法については後述するが、例えば、(1)塩基配列長、(2)融解温度、(3)2次構造の生成可能性、(4)塩基配列の類似性を利用して選別される。110は最終的に得られた最適プローブを示す。なお、

図1に示すプローブ設計処理において、入力ターゲット塩基配列107であり、出力は最適プローブ110である。

#### 【0016】

図1の各データ及び処理について詳述する前に、本発明のバックグラウンドとなる知識を説明する。図3はDNAマイクロアレイ上のハイブリダイゼーション（以降、ハイブリと略す場合がある）の様子を示した図である。生体内でほとんどの場合、DNA塩基列は2重らせん構造をしていて、その2本鎖の間の結合は塩基間の水素結合で実現されている。一方、RNA塩基列は1本で存在する場合が多い。塩基の種類はDNAの場合はACGTの4種類、RNAの場合はACGUの4種類であり、それぞれ水素結合ができる塩基対はA-T（U）、G-Cのペアとなっている。ハイブリダイゼーションとは、1本鎖状態の核酸分子同士がある部分の塩基配列を介して部分的に結合する状態をいい、本実施形態で想定している反応は、図3の上側の基板にくっついた核酸分子（プローブ）の方が下側の資料中にある核酸分子より短い。よって、資料中に存在する核酸分子がプローブ塩基配列を含む場合は、このハイブリダイゼーション反応はうまくいき、試料中のターゲット核酸分子はトラップされることとなる。

#### 【0017】

次に、図4を用いて感染症の菌を特定するDNAマイクロアレイの原理を説明する。図4で示したDNAマイクロアレイは、例えば、黄色ブドウ球菌を特定する目的で作られていると仮定する。図4の左側は、黄色ブドウ球菌野生株由来の処理系列であり、右側は大腸菌野生株由来の処理系列である。例えば、左側は黄色ブドウ球菌に感染した患者の血液を処理する流れで、右側は大腸菌に感染した患者の血液を処理する流れだと考えてよい。

#### 【0018】

どちらも基本的には同じ処理を行う。つまり、まず初めに例えば菌感染患者の血液や、痰などからDNAを抽出する（401、411）。この際に、一般的には、患者の体細胞由来の人間のDNAも含まれる可能性がある。抽出されたDNAが少ない場合、PCRなどの方法で増幅を行う。この際に蛍光物質もしくは蛍光物質を結合させることができる物質を標識として混入させるのが一般的である（402、412）。

#### 【0019】

増幅をしない場合は、抽出されたDNAを用いて、相補鎖を作りながら蛍光物質もしくは蛍光物質を結合させることができる物質を標識として混入させる、または、そのまま直接抽出されたDNAに蛍光物質もしくは蛍光物質を結合させることができる物質を標識として付加させる（403、413）。

#### 【0020】

通常、PCR増幅を行う場合、感染症の菌特定が目的であれば、いわゆる16sといわれるリボゾームRNA（16s rRNA）を構成する塩基配列の部分を増幅するのが一般的である。この場合、図4の左側の黄色ブドウ球菌のPCRプライマーと右側の大腸菌のPCRプ

ライマーはほとんど同じものを使うこととなる。より具体的には、どんな菌の16s rRNAをコーディングしている部分でも増幅させることができるプライマーセットを用いて、マルチプレックスPCRを行うことが好ましい。この場合、結果的に、図4における右と左のどちらのハイブリ溶液(404、414)も複数の種類の塩基配列を含むものとなる。この理由は図5を参照して詳しく述べる。

#### 【0021】

これに対して、より詳しい塩基配列解析を行いたい場合は、例えば、黄色ブドウ球菌用のPCRプライマーセット、大腸菌用のPCRプライマーセットを別々に設定する。この場合、仮に菌のゲノムの特定部分だけを選択的に増幅するようにプライマーを設定すれば、ハイブリ溶液に含まれる塩基配列の種類は非常に限定される。それでも通常、自然界に存在する菌の株は数種類に及ぶので、ハイブリ溶液に存在する塩基配列の種類が1種類となることは稀である。

#### 【0022】

ここで黄色ブドウ球菌を判定する目的のために設計されたDNAマイクロアレイが正しく動作するならば、ハイブリ溶液404ではスポットがポジティブに反応し(405)、右側のハイブリ溶液414では、スポットがネガティブに反応する(415)。

#### 【0023】

これと全く同じように、大腸菌の存在を判定する目的のために設計されたDNAマイクロアレイが正しく動作するならば、ハイブリ溶液404ではスポットがネガティブに反応し、ハイブリ溶液414ではスポットがポジティブに反応する。もちろん、いろんな菌に対してそれぞれ特異的に反応する数種類のスポットを同時に並べたDNAマイクロアレイを用いて、感染菌の判定を行ってもかわまない。

#### 【0024】

次に、図5を用いて、図4のハイブリ溶液に複数の種類の塩基配列が存在する理由を説明する。通常、自然界に存在する菌は、突然変異を頻繁に起こす。その結果、淘汰を経て生き残った主要な数種類の株が同時に存在することがある。例えば、院内感染などで問題を起こす菌株は、通常は薬剤耐性がないはずの菌が、突然変異を起こすことによって薬剤耐性を獲得することで出現する。薬剤耐性を獲得した結果、殺菌努力を行っている衛生的な環境でも旺盛な繁殖力を持つ菌が出現したりする。このように、自然界に存在する同一の菌の塩基配列は、数種類のバリエーションをもつと考えるのが妥当である。

#### 【0025】

図5に示したのは、黄色ブドウ球菌にMu50とMW2の2種類の株のゲノム構造である。それぞれの株のゲノムの総塩基数はそれぞれ2,878,040と2,820,462と異なっている。なお、図5では、左から右への方向を5'末端から3'末端方向とし、これを順方向として塩基配列が示されている。更に16sのリボゾームRNA(16s rRNA)をコーディングしている部位は、Mu50の場合は順方向2箇所、逆方向3箇所の合計5箇所であるのに対して、MW2では順方向3箇所、逆方向3箇所の合計6箇所存在する。これらの16s rRNAの各部位の塩基配列はそれぞれ非常に似ているが、全く同じではない。つまり、調べようとしている菌感染患者の体の中に存在する菌の株の種類がたとえ1種類であったとしても、図4のような一般的な処理でハイブリ溶液を調製すると複数種類の塩基配列がハイブリ溶液の中に存在することになる。この複数の塩基配列に対して同じようなハイブリ強度を示すプローブを設計することが本実施形態のプローブ設計方法の目的である。

#### 【0026】

この目的を達成するために本実施形態のプローブ設計方法は、図1に示した通り、ターゲット塩基配列と同じグループに属する塩基配列のセット(自己塩基配列データ101)と、そのグループに競合するグループに属する塩基配列のセット(競合塩基配列データ103)から別々に頻度テーブルを作成する。図4の例で言うと、例えば黄色ブドウ球菌の色々な株の色々な部位の16s rRNAの塩基配列の集合が自己塩基配列データ101となり、黄色ブドウ球菌以外の例えば大腸菌や、インフルエンザ菌などの種々の菌の色々な株の色々な部位の16s rRNAの塩基配列の集合が競合塩基配列データ103となる。

## 【0027】

これらの塩基配列データから頻度テーブルを作成する様子を図6に示す。頻度テーブルの作成においては、塩基配列データに含まれる長さ  $n$  (図6では  $n=9$ ) の部分配列の出現回数をカウントする。長さ  $n$  の塩基配列のバリエーションは4の  $n$  乗なので、図6において、行の数は  $4^n$  となる。なお、図6において出現頻度が低いほど、その部分配列の特異性は高くなるので、例えば頻度にマイナス1をかけた数を特異性とする。

## 【0028】

すなわち、自己頻度テーブル作成ステップ102は検出対象の菌種の種々の株の16s rRNAの塩基配列を記憶する自己塩基配列データ101を参照して、長さ  $n$  の全ての部分塩基配列について出現回数をカウントし、その結果を図6の如くテーブル化し、自己頻度テーブル105とする。同様に、競合頻度テーブル作成ステップ104は検出対象の菌種以外の菌種(検出対象から区別されるべき菌種)の種々の株の16s rRNAの塩基配列を記憶する競合塩基配列データ101を参照して、長さ  $n$  の全ての部分塩基配列について出現回数をカウントし、その結果を図6の如くテーブル化し、競合頻度テーブル106とする。

## 【0029】

次に、図1のターゲット塩基配列107に対する部分列( $n$ 個の塩基からなる塩基配列)の頻度または特異性を上記頻度テーブルを用いて求める。この様子を示したのが図7である。このターゲット塩基配列107は、例えば、図5におけるMu50株の黄色ブドウ球菌の一番左の16s rRNAコーディング部位(配列X)などとなる。プローブ評価ステップ108では、自己頻度テーブル105及び競合頻度テーブル106を参照して、ターゲット塩基配列107から順次得られる部分列の頻度(特異性)を取得し、評価する。この様子をグラフに示したのが図8である。

## 【0030】

図8の横軸はターゲット塩基配列の部分配列の位置であり、例えば16s rRNAの部分を調べたいとすれば、その長さは約1500程度となる。縦軸は、その位置における部分配列の特異性であり、例えば頻度テーブルの値にマイナス1をかけた数である。図8において、上段に示したグラフが図1における競合頻度テーブル106を元に導いた特異性を表し、下段に示したグラフが図1における自己頻度テーブル105を元に導いた特異性を表す。

## 【0031】

例えば、図8のグラフの場合、配列の先頭から約2/3の部分、つまり約1000ベースぐらいの位置で、このターゲット塩基配列の特異性の高い部分が存在する。例えば、このターゲット塩基配列が図5のMu50株の最初の16s部分(配列X)だったとすると、この配列Xの1000ベースあたりの位置で、この配列に特有の塩基配列が存在し、それは、Mu50株のその他の16sの部位、または、黄色ブドウ球菌のMu50株以外の株の16s rRNA部位とは異なる配列であることを意味する。よって、この部分からプローブ候補を選択することは、好ましくない。また、上段のグラフにおいて特異性の低い部分は、他の種の菌との識別ができないことを意味しており、その部分からプローブ候補を選択することも好ましくない。

## 【0032】

したがって、上段のグラフで山となっており(異種菌間で特異性が高い)、かつ下段のグラフで山となっていない部分(同種菌の異種株間で特異性が低い)よりプローブ候補を選択する。このことで、判定すべき菌とは異なる菌種に対しては特異性が高く(=頻度が低く)、判定すべき菌の色々な変異体に対しては特異性が低い(頻度が高い)プローブを選ぶことができる。すなわち、図8のように、各特異性を複数の部分塩基配列を基準に対比可能に表示することにより、ユーザは容易に適切なプローブ候補を選択できる。

## 【0033】

この結果、同じ菌であれば、どのような株のどのような部位の16s rRNAコーディング配列がハイブリ溶液に含まれていても強いハイブリ反応を示し、かつ、他の菌のどのような株のどのような部位の16s rRNAコーディング配列がハイブリ溶液に含まれていても弱いハ

イブリ反応を示すプローブが選ばれることとなる。

#### 【0034】

なお、本発明のプローブ設計方法は、感染症の判別にのみ適用できるものではなく、一般的に同一と判断すべき塩基配列の対象の中に若干のバリエーションが存在するものなら何でも適用できる。例えば人間の個体判別によく用いられるMHCなどの判定などにも適用できる。

#### 【0035】

次に、図1のプローブセット選択ステップ109について説明する。図1のプローブ評価ステップ108で高い評価を得たプローブの中で、長さを揃えてプローブセットとするのが一番単純なプローブセット選択の方法である。しかしながら、ハイブリダイゼーション反応はプローブ塩基配列の長さよりも、その融解温度で決まる場合が一般的である。よって、例えば標準プローブ長 $n$ （本例では $n=24$ とする）を設定しておいて、その長さの前後数塩基の許容範囲（本例では $\pm 2$ とする）より得られる各プローブについて融解温度を求め、出来るだけ融解温度を一定するように採用すべきプローブを決定することによって、より高い品質のプローブセットが得られる。

#### 【0036】

なお、塩基配列の融解温度は、その配列の塩基の混入率で求める方法や、いわゆるNearest Neighbor法と呼ばれる連続する2つの塩基配列の並び方から求める方法などが知られており、これらの方法を用いて融解温度を算出することができる。

#### 【0037】

また、塩基配列の長さが20を超えてくると、2次構造を形成して、プローブとして適切でない場合が発生する。これを防ぐために、例えばMichael Zuker氏の考案した“Calculating nucleic acid secondary structure.”（Current Opinion in Structural Biology, 10, 303-310, (2000)）に記載の方法などを用いて2次構造形成の可能性を計算し、2次構造を形成しやすいプローブは排除するようにしてもよい。

#### 【0038】

また、図1のプローブ評価ステップで説明した頻度テーブルを用いる方法は、あくまで特異性の目安としてプローブ候補を選ぶ方法であるので、実際に選ばれたプローブ群が似かよったものとなる可能性がある。したがって、プローブ候補がどの程度お互いに似ているかを、お互いのマッチングを取ることで調べ、似かよったプローブ候補を排除することが好ましい。例えば、 $N$ 個のプローブ候補からプローブセットを選択するとすると、 $N(N-1)/2$ 個のマッチングを取り、プローブ同士の塩基配列がお互いにどれぐらい異なっているかを評価して、異なる塩基数がより多いプローブセットを選ぶようにする。このことによって、より高い品質のプローブセットを選択することができる。なお、この方法は、いわゆるクロスハイブリダイゼーションを防ぐ方法と呼ばれる。

#### 【0039】

[プローブ設計装置のより詳細な説明]

図9～図11を用いて第1実施形態のプローブ設計プログラムの動きを説明する。プローブ設計プログラムの動作はまずターゲット生物グループの選定（901）から始まる。例えば、感染症の原因菌特定のためのプローブを設計しようとしたとすると、ターゲット生物グループが、菌やウイルス、真菌などのゲノム情報が塩基配列データベース906から選ばれる。図9において、塩基配列データベース906は、NCBIなどの公共のデータベースや、自社のイントラネット上に構築された塩基配列データベースなどで、現在入手可能な最大量のデータが蓄積されたものを意味する。一方、ターゲット塩基配列データベース907は、ターゲット生物グループの選定（901）で選ばれた生物種のゲノム情報しか含まない。例えば、本プログラムを人間の体質判定のためのプローブ設計に適用した場合は、ターゲット塩基配列データベース907に格納される塩基配列は、例えばMHCのDRB1の部分の全アレルの情報などとなる。

#### 【0040】

次に、ターゲット生物種の選定を行う（902）。ターゲット生物種を選定すると、タ

ターゲット塩基配列データベース 9 0 7 に含まれる塩基配列が自己塩基配列データと競合塩基配列データに分けられる。すなわち、ターゲット塩基配列データベース 9 0 7 に含まれる情報から、ターゲット生物種の複数の多型、複数のゲノム部位に対応する自己塩基配列データ 9 0 8 が抽出されるとともに、ターゲット塩基配列データベース 9 0 7 に含まれる、ターゲット生物種以外の生物種の複数の多型、複数のゲノム部位に対応する競合塩基配列データ 9 0 9 が抽出される。

#### 【0 0 4 1】

次に、選ばれたターゲット生物種を元に自己頻度テーブル 9 1 0 (図 1 の自己頻度テーブル 1 0 5 に相当) と競合頻度テーブル 9 1 1 (図 1 の競合頻度テーブル 1 0 6 に相当) を作成する (9 0 3)。この頻度テーブルの作成 (9 0 3) のとき、通常は、ターゲットとするゲノム領域も設定する。例えば菌種の判別を行うプローブを設計する場合には、16 s rRNA の部分が選ばれたりする。

#### 【0 0 4 2】

より具体的に述べると、図 4 で述べたとおり、通常、DNA マイクロアレイを用いた実験を行う場合、PCR でターゲットとなる核酸を増幅する。この時、PCR プライマーで挟んだ領域のみが増幅されるので、ターゲット塩基配列データベース 9 0 7 の中の PCR で増幅される部分のみを用いて頻度テーブルを作成する。本プログラムを人間の体質判定のためのプローブ設計に適用した場合は、例えばターゲットとする DNA 領域も MHC の DRB 1 の部分などに予め設定されるので、この DNA 領域を設定する必要はない。また、例えば MHC の DRB 1 の解析であれば、たんぱく質発現に差を示さない 3 桁までのアレルを同一種として扱う。なお、ターゲットとする核酸領域、または PCR 増幅領域の指定は、通常、プログラムのプロパティとして設定し、UI (ユーザインターフェース) を使ってその都度設定することはない。

#### 【0 0 4 3】

次にターゲット塩基配列の選定を行い (9 0 4)、そのターゲット塩基配列の部分列群の特異性を自己頻度テーブル 9 1 0 と競合頻度テーブル 9 1 1 を用いて評価し、プローブを選定する (9 0 5)。

#### 【0 0 4 4】

図 1 0 は、ターゲット生物種の選定 (9 0 2) からターゲット塩基配列の選定 (9 0 4) までを行うユーザインターフェースの例を示したものである。まず、ターゲット菌種欄 1 0 0 1 にターゲット菌種としての菌の一覧が表示される。ここで、例えば黄色ブドウ球菌 (*Staphylococcus aureus*) を選ぶと、株表示欄 1 0 0 2 に黄色ブドウ球菌の色々な株の色々な部位の 16s rRNA を示す一覧が表示される。ここまでのインタラクションで、図 9 の 9 0 2、9 0 3 (ターゲット生物種の選定と頻度テーブルの作成) が実行される。株表示欄 1 0 0 2 に示されている多型としての複数の株は、図 5 に示した株である。例えば、図 1 0 のように、株の名前、ゲノム上の位置、方向を示すことによって、それぞれの塩基配列の識別ができるようにする。

#### 【0 0 4 5】

株表示欄 1 0 0 2 の一覧の中で 1 つの配列を選ぶと、ターゲット塩基配列の選定 (9 0 4) が実行される。図 1 0 に示したように黄色ブドウ球菌 Mu50 株の先頭の 16s rRNA を選んだとすると、ターゲット塩基配列として図 5 の配列 X が選ばれたことになる。そして、デザインボタン 1 0 0 3 を押すと、プローブ評価ステップ 1 0 8 に処理が移行し、デザイン画面が表示される。

#### 【0 0 4 6】

図 1 1 は実際のデザイン画面の例を示した図である。1 1 0 1 から 1 1 0 4 はグラフを示し、横軸は、例えば図 1 0 のような UI で選ばれたターゲット塩基配列の位置を意味し、それぞれのグラフはそのターゲット塩基配列のそれぞれの位置の部分塩基配列に対する値をグラフ化したものである。グラフ 1 1 0 2 とグラフ 1 1 0 3 は、それぞれ図 8 の上段下段のグラフに対応しており、グラフ 1 1 0 2 はそれぞれの位置の部分配列の競合塩基配列データに対する特異性をグラフ化したもの、グラフ 1 1 0 3 はそれぞれの位置の部分配

列の自己塩基配列データに対する特異性をグラフ化したものとなる。また、グラフ 1101 はそれぞれの位置の部分配列の人間のゲノムに対する特異性を示す。グラフ 1104 はそれぞれの位置の部分配列を先頭とする所定個数の塩基配列（本例では 24 個の塩基からなる塩基配列（24 ベースの塩基配列））の融解温度を示す。

#### 【0047】

図 8 で説明したように、グラフ 1102 で山になっていて、グラフ 1103 で谷になっている場所でプローブを設定するのがよい。1105 は情報欄であり、現在のターゲット生物種や各種のパラメータなどを表示する。なお、本実施形態ではデフォルトの塩基長として 24 が設定されており、これを基準にグラフ 1104 の融解温度が計算されている。1106 は、設計したプローブのリストで、その位置は、ライン 1107（点線）で表示される。1108 の実線は“カレント”位置を示し、現在注目している位置で、その位置に対応する部分塩基配列（24 ベースの塩基配列）は 1109 に表示され、その直前の部分の塩基配列は 1110、その直後の部分の塩基配列は 1111 に表示される。本実施形態では前後 10 個の塩基からなる配列が示される。

#### 【0048】

また、図 11 の UI では、1112 に示すように、ターゲット塩基配列の中で塩基配列を検索する機能を持っている。

#### 【0049】

また、例えば、感染症の原因菌を特定するプローブを設計する場合は、人間のゲノムも含まれてくるので、例えばグラフ 1101 に示すように、人間のゲノムに対するターゲット塩基配列の部分配列の特異性を表示してもよい。また、プローブを設定するために実施例 1 で説明したように融解温度（ $T_m$ ）を選択したプローブ間で出来るだけ似通った温度にする方がいい。これを実現するために、例えばグラフ 1104 では、その位置の部分塩基配列の  $T_m$  をグラフとして表示する。

#### 【0050】

プローブ評価ステップ 108 では、グラフ 1101～1104 を参照してユーザが指示した実線 1108 の移動に応じてプローブ候補の評価を行なう。このとき、例えば、実線 1108 を移動させていく間に指定可能な位置（グラフ 1103 において特異性がある第 1 の閾値よりも低く、グラフ 1102 において特異性がある第 2 の閾値よりも高い部分）に入った際にこれをユーザに報知する（例えば、実線 1108 の色を変える等）ように構成する。このようにすれば、ユーザはより容易に適切な部分塩基配列を見つけることができる。そして、実線 1108 が上記指定可能な位置にある間に、不図示の確定ボタンを押すことにより、当該位置に対応した部分塩基配列がプローブ候補として設定される。こうして設定されたプローブ候補はさらにプローブセット選択ステップ 109 でさらに絞り込まれ、適切なプローブセットが決定される。

#### 【0051】

なお、上記プローブ評価ステップ 108 において、グラフ 1103 で特異性が低く、グラフ 1102 で特異性が高い部分を自動的に抽出してユーザに提示するようにしてもよい。例えば、グラフ 1103 において特異性がある第 1 の閾値よりも低く、グラフ 1102 において特異性がある第 2 の閾値よりも高い部分を抽出してユーザに提示する。なお、上記で設計された塩基配列の相補的な配列も同様にプローブとして用いることができるので、これらを併せて表示する、或いは設計結果として提示するようにしてもよい。

#### 【0052】

##### <第 2 実施形態>

第 1 実施形態では、図 11 のように頻度情報を表示して、ユーザに適切な位置を選択させている。このように頻度情報を用いることで、ユーザは容易に適切なプローブ候補を選択できるのであるが、一般にマイクロアレイに設定するプローブの数は数百から数千と膨大である。よって、このような全てのプローブを頻度情報を元にしてユーザが設定していくのは、大変な時間と手間がかかることにもなりうる。また、第 1 実施形態で述べたように、単純に特異性の値を閾値と比較することにより自動的に部分塩基配列を抽出すること



もできる。しかしながら、この場合、(1) 塩基長の全体に渡ってサーチとは多大な計算時間が必要となる、(2) 類似した塩基配列が数多く抽出される可能性がある、(3) よって、塩基長の全体に渡って、適度に分散した位置から部分塩基配列を抽出することが困難である、といった課題がある。

#### 【0053】

そこで、第2実施形態においては、上記の課題を解決するプローブ設計の自動化方法について説明する。

#### 【0054】

第2実施形態によるプローブ設計方法が適用される情報処理装置の構成は第1実施形態(図2)と同様である。

#### 【0055】

図12は、第2実施形態によるプローブ設計方法の処理手順を説明するフローチャートである。第1実施形態(図1)と同様のステップ及びデータには同一の参照番号を付してある。

#### 【0056】

1201は全塩基配列データであり、実際には自己塩基配列データ101と競合塩基配列データ103の集合である。1202は共通配列データ作成ステップであり、全塩基配列データ1201に含まれる全ての塩基配列において共通する部分塩基配列を抽出し、共通領域データ1203を作成する。なお、共通する部分塩基配列は所定数以上の塩基からなる塩基配列(例えば20ベース以上の塩基長)であって、全ての塩基配列中に含まれるものを検索することにより得られたものとする。また、このよう共通領域は、例えば、マルチプルアライメント処理により得ることができる。

#### 【0057】

1211はプローブ評価ステップであり、自己頻度テーブル105と競合頻度テーブル106と共通領域データ1203を用いて、ターゲット塩基配列107からプローブ候補を導き出すステップである。プローブセット選択ステップ1212は、プローブ評価ステップ1211の結果得られたプローブ候補より最適なプローブセットを選択する。1213は最終的に得られた最適プローブを示す。なお、図12に示すプローブ設計処理において、入力ターゲット塩基配列107であり、出力は最適プローブ1213である。

#### 【0058】

ここで、本実施形態による自動プローブ設計について概要を説明する。プローブ設計の自動化のために、本実施形態では共通領域データ1203を利用する。共通領域データ1203は、共通領域データ作成ステップ1202により作成される。

#### 【0059】

共通領域データ作成ステップ1202では、自己塩基配列データ101と競合塩基配列データ103を合わせた全塩基配列データ1201を対象に、全塩基配列において共通に存在する部分配列を探し、その配列における位置と長さを共通領域データ1203として保持する。例えば、細菌の16s rRNAにおいては、共通に存在する部分配列が似通った位置に存在することが分かっている。

#### 【0060】

ターゲット塩基配列107について共通領域データ1203を参照すると、図13のように、ターゲット塩基配列を意味する1301上では、共通領域1302、1303…1306…、とそれらの共通領域に挟まれた領域1312、1313、…、1315…とを区別することが可能になる。この共通領域に挟まれた各領域において、自己頻度テーブル105と競合頻度テーブル106とを参照することで、同種菌の異種株間での特異性が低くかつ異種菌間で特異性の高くなる位置を一箇所選択する。ターゲット塩基配列1301上には、共通領域は複数箇所あり、よって共通領域に挟まれている領域も複数存在するため、共通領域に挟まれている領域のある限り、同様の処理を機械的に繰り返すことで、ターゲット塩基配列1301の全長に渡ってプローブを設定することが可能となる。以上の処理については、複数の菌種に渡って多数のターゲット塩基配列107がある場合でも機



械的に処理することが可能であり、よってコンピュータによる自動化が可能である。

#### 【0061】

なお、本実施形態のプロブ設計方法は、第1実施形態と同様に、感染症の判別にのみ適用できるものではなく、一般的に同一と判断すべき塩基配列の対象の中に若干のバリエーションが存在するものなら何でも適用できる。例えば人間の個体判別によく用いられるMHCなどの判定などにも適用できる。

#### 【0062】

また、共通領域データ作成ステップ106を通して作成された共通領域データ109は、多種の遺伝子を共通して増幅させることができるPCR用のユニバーサルプライマとしての利用も可能である。

#### 【0063】

[プロブ設計装置のより詳細な説明]

図14～図16を用いて第2実施形態のプロブ設計プログラムの動きを説明する。14において第1実施形態(図9)と同様の処理及びデータについては同一の参照番号を付してある。図9に関して上述したように、ターゲット生物グループの選定(901)では、設計しようとするプロブに応じて選定されたターゲット生物グループに属する、菌やウイルス、真菌などのゲノム情報が塩基配列データベース906から選ばれ、ターゲット塩基配列データベース907に格納される。

#### 【0064】

次に、ターゲット生物種の選定(902)により、ターゲット塩基配列データベース907に含まれる塩基配列が自己塩基配列データと競合塩基配列データに分けられる。すなわち、ターゲット塩基配列データベース907に含まれる情報から、ターゲット生物種の複数の多型、複数のゲノム部位に対応する自己塩基配列データ908が抽出されるとともに、ターゲット塩基配列データベース907に含まれる、ターゲット生物種以外の生物種の複数の多型、複数のゲノム部位に対応する競合塩基配列データ909が抽出される。

#### 【0065】

次に、選ばれたターゲット生物種を元に自己頻度テーブル910と競合頻度テーブル911を作成する(903)。この頻度テーブルの作成(903)のとき、通常は、ターゲットとするゲノム領域も設定する。例えば菌種の判別を行うプロブを設計する場合には、16s rRNAの部分が選ばれたりする。

#### 【0066】

頻度テーブルの作成(903)とともに、共通領域データ913(図12の共通領域データ1203に相当)を作成する(921)。即ち、ターゲット塩基配列データベース907に保持されている全ての塩基配列(912)に共通する部分配列の情報が共通領域データ913に格納される。頻度テーブルの作成については第1実施形態で詳述したとおりである。次にターゲット塩基配列の選定を行い(922)、そのターゲット塩基配列の部分列群の特異性を自己頻度テーブル910と競合頻度テーブル911と共通領域データ913を用いて評価し、プロブを選定する(923)。

#### 【0067】

図15は、ターゲット生物種の選定(902)からターゲット塩基配列の選定(922)までを行なうユーザインタフェースの例を示したものである。まず、ターゲット菌種欄1501にターゲット菌種としての菌の一覧が表示される。ここで、例えば黄色ブドウ球菌(*Staphylococcus aureus*)を選ぶと、塩基配列の表示欄である1502に黄色ブドウ球菌の種々の株の色々な部位の16S rRNAの塩基配列が一覧として表示される。ユーザはこの1502の表示欄から所望の株の所望の部位を選択できる。ここまでのインタラクションで、図14の902、903、921(ターゲット生物種の選定と頻度テーブル及び共通領域データの作成)が実行される。

#### 【0068】

例えば黄色ブドウ球菌であっても、図5に示すように、株ごとに複数の部位に16S rRNA領域があることから、任意の株における任意の位置の16S rRNA領域を選択可能にするため

に、表示欄 1502 にて選択された塩基配列に関する情報を表示欄 1503 に表示する。表示欄 1503 のように株名称、ゲノム上の位置の情報などを示すことにより、それぞれの塩基配列の識別が可能である。また、表示欄 1503 には、GI や ACCESSION 番号のような公共データベースにおいて一意に定められている識別番号を表示したり、同時にその識別番号から公共データベースの情報を表示できるような仕組みを備えても良い。

#### 【0069】

塩基配列表示欄である 1502 は、複数の塩基配列の位置を調製して（例えば、マルチプルアライメント処理）して、表示しても良い。また、塩基配列表示欄である 1502 は、各塩基配列における共通領域に相当する部分を、色やフォントなどを変更することによって強調表示してもよい。

#### 【0070】

塩基配列を表示する表示欄 1502 の一覧の中で一つの配列を選ぶと、ターゲット塩基配列の選定（922）が実行される。図 15 に示したように黄色ブドウ球菌 Mu50 株の先頭の 16s rRNA を選んだとすると、ターゲット塩基配列として図 5 の配列 X が選ばれたことになる。そして、デザインボタン 1004 を押すと、図 16 に示すようなデザイン画面が表示される。

#### 【0071】

図 16 は本実施形態によるデザイン画面の例を示した図である。インターフェースの構成は第 1 実施形態（図 11）とはほぼ同様であり、同様の構成要素には同一の参照番号を付してある。

#### 【0072】

1101 から 1104 はグラフを示し、横軸は、例えば図 15 のような UI で選ばれたターゲット塩基配列の位置を意味する。図 11 で上述したように、グラフ 1102 はそれぞれの位置の部分配列の競合塩基配列データに対する特異性をグラフ化したもの、グラフ 1103 はそれぞれの位置の部分配列の自己塩基配列データに対する特異性をグラフ化したものである。また、グラフ 1101 はそれぞれの位置の部分配列の人間のゲノムに対する特異性を示す。グラフ 1104 はそれぞれの位置の部分配列を先頭とする所定個数の塩基配列（本例では 24 個の塩基からなる塩基配列（24 ベースの塩基配列））の融解温度を示す。

#### 【0073】

先に記したように、図 15 のような UI にてターゲット塩基配列を選択した時点で、自己頻度テーブル 910 と競合頻度テーブル 911 と共通領域データ 913 を用いて、プローブの自動設計を行なうことも可能である。その場合は、図 16 のデザイン画面には既に作成されているプローブが表示されることとなる。また、作業者によりプローブ位置の編集（追加・削除も含む）も可能である。

#### 【0074】

グラフ 1101、1102、1103 に記されているポイント 1601 は、図 14 における共通領域データ 912 より導かれる共通領域データ 913 の位置を示している。共通領域の位置を示すポイント 1114 に挟まれているそれぞれの領域内において、グラフ 1102 で山になっていてグラフ 1103 で谷になっている場所を自動的に選択していくことにより、プローブが自動で設定される。なお、ユーザが操作して手動でプローブを設定（部分塩基配列を指定）するマニュアルモードを設けてもよいことは明らかである。このマニュアルモードにおいては、第 1 実施形態（図 8 或いは図 11）でも説明したように、グラフ 1102 で山になっていて、グラフ 1103 で谷になっている場所でプローブを設定するのがよい。なお、この場合、ポイント 1601 として表示されている共通領域では部分塩基配列の先頭位置を指定できないように構成してもよい。

#### 【0075】

1105 は情報欄であり、現在のターゲット生物種や各種のパラメータなどを表示する。なお、本実施形態ではデフォルトの塩基長として 24 が設定されており、これを基準にグラフ 1104 の融解温度が計算されている。1106 は、設計したプローブのリストで

、その位置は、ライン1107（点線）で表示される。1108の実線は“カレント”位置を示し、現在注目している位置で、その位置に対応する部分塩基配列（24ベースの塩基配列）は1109に表示され、その直前の部分の塩基配列は1110、その直後の部分の塩基配列は1111に表示される。本実施形態では前後10個の塩基からなる配列が表示される。

#### 【0076】

なお、上記で設計された塩基配列の相補的な配列も同様にプローブとして用いることができるので、これらを併せて表示する、或いは設計結果として提示するようにしてもよいことはいふまでもない。

#### 【0077】

サーチボタン1112が押されると、ターゲット塩基配列よりプローブに適した部分塩基配列を抽出するべく、探索を開始する。上述しているように、第2実施形態では、自己頻度情報・競合頻度情報・共通領域情報を用いて部分塩基配列を検索し、プローブの自動設計を行う。以下、本実施形態のプローブ自動設計方法について図17を用いて、さらに詳細に説明する。図17は、感染症起炎菌用のプローブを、プローブ長を $24 \pm 2$ 、融解温度を $50 \pm 1$ 度という条件で作成する例である。

#### 【0078】

ターゲット塩基配列1401上に、1402～1406までの5箇所の共通領域があるとした場合、まず共通領域1402と1403の間の領域1407でプローブ作成を試みる。共通領域1402と1403の間の領域1407においてプローブ作成に成功しようとしまいと、次の共通領域間の領域（共通領域1403と1404の間の領域）で同様にプローブ作成を試みる。そして、その次は共通領域1404と1405の間で、さらにその次は共通領域1405と1406の間で、というように順番に全ての共通領域に挟まれている間の領域でのプローブ作成を試みる。

#### 【0079】

プローブ作成の試みの手順は次のとおりである。領域1407内で、競合頻度グラフ1408から得られる他の菌種の塩基配列に対する特異性を示す指標が高く、かつ、自己頻度グラフ1409から得られる同一の菌種の塩基配列に対する特異性を示す指標が低い位置1410を抽出する。そして、その位置1410を基準に融解温度を調べるための候補となる部分配列群1411を作成する。本実施形態では、他の菌種の塩基配列に対する特異性を示す指標が最大の位置を検出し、この位置において同一の菌種の塩基配列に対する特異性を示す指標が所定値よりも低い場合に、候補位置の抽出に成功したものとする。候補位置の抽出に失敗した場合は、次の領域へ処理を移す。そして、ターゲット塩基配列1401より、候補位置1410を先頭に24塩基長の部分配列を取り出す。

#### 【0080】

次に、候補配列作成1411において、候補位置1410を先頭に取り出した24塩基長の部分配列に対して、先頭或いは末尾或いは両方に1～2塩基を追加したり、削除したりすることにより、開始位置、終了位置、塩基長にバリエーションを持つ複数の候補部分塩基配列を作成する。そして、1412において、1411で得られた複数の候補配列のそれぞれについて融解温度を計算する。そして、1413において、想定している温度条件（本例では $50 \pm 1$ 度）の範囲内にあり、かつ、想定している温度（本例では50度）に最も近い融解温度を持つ配列を抽出する。このようにして、共通領域1402と共通領域1403の間の領域1407において、プローブ1414を求めることができる。なお、融解温度を計算した結果、1411で得られた複数の部分配列の全てについて、設定された条件に合う融解温度とならないことが判明した場合は、共通領域1402と共通領域1403の間にプローブを作成することをあきらめ、次の共通領域1403と共通領域1404に挟まれた領域におけるプローブ作成を試みる。

#### 【0081】

以上、図17にて説明したプローブの自動設計処理について、図18のフローチャートを参照してさらに説明する。

**【0082】**

図18は、ターゲット遺伝子の核酸配列を選択後に、その核酸配列からプローブを自動的に選択する手順を示すフローチャートである。図15に示したフローチャートを、感染症の原因菌の16S rRNAを対象に、24塩基±2塩基の長さで、融解温度を50度±1度と設定する例により、以下、説明する。

**【0083】**

まず、ターゲット遺伝子の核酸配列の5'側末端から3'側末端に向けて共通領域を探す(ステップS1501)。共通領域が存在すれば(ステップS1502)、ターゲット遺伝子の5'側末端から共通領域までの間を、1塩基ずつずらしながら競合頻度テーブルを参照し、競合率に対して最も特異性の高くなる部分配列の位置を見つけた(ステップS1503)。次に、この位置について自己頻度テーブルを参照し同種の塩基配列に対する特異性を調べ(ステップS1504)、同種の塩基配列に対する特異性が十分に低いと判断される値であるかどうかを判断する(ステップS1505)。十分に低いと判断する基準は、例えば自己頻度テーブルの各値の平均値より低いかどうか、あるいは、あらかじめ規定する任意の閾値より低いかどうか、等、状況に応じて決めておく。

**【0084】**

同一種に対しても特異性の高い部分であると判断される場合は、ステップS1501に戻り、現在の位置の次にある共通領域より3'側に共通領域があるかどうかを確認する。ステップS1505にて、自己頻度テーブルから得られる特異性の値が十分に低いと判断された場合は、現在位置を基準に前後に塩基を追加或いは削除して22~26塩基長となる部分配列を作成し(図17の1411)、作成されたそれぞれの部分配列について融解温度を計算する(ステップS1506)。そして、融解温度が50度に最も近くなる塩基配列を選び、その融解温度が49度から51度の間であればその塩基配列をプローブとする(ステップS1507~S1509)。一方、選択した塩基配列の融解温度が49度~51度の範囲に無い場合は、ステップS1501に戻り、現在の位置の次にある共通領域の、より3'側に共通領域があるかどうかを確認する。

**【0085】**

このように、ターゲット塩基配列全体に対して、共通領域で区切った部分ごとに、その中での特異性を計算し、融解温度を計算することで、ターゲット塩基配列全体に分布するプローブセットを作成することができる。

**【0086】**

[プローブセットの設計例及び実験例]

次に図19を用いて、上記実施形態のプローブ設計方法を用いて設計されたプローブを用いたDNAマイクロアレイの実験手順について説明する。

**【0087】**

1201の“サンプル”とは対象としている核酸が含まれているはずの液体や個体である。例えば感染症の原因菌の特定をする場合、ヒト、家畜等の動物由来の血液、喀痰、胃液、膣分泌物、口腔内粘液等の体液、尿及び糞便のような排出物等細菌が存在すると思われるあらゆる物がサンプルとなる。また、食中毒、汚染の対象となる食品、飲料水及び温泉水のような環境中の水、空気清浄機や浄水器等のフィルタ等、細菌による汚染が引き起こされる可能性のある媒体がサンプルとして用いられることもある。さらに、輸出入時における検疫等の動植物も検体としてその対象となる。

**【0088】**

次に、“生化学的増幅”方法を用いてサンプル1201を増幅する(1202)。例えば感染症の原因菌の特定をする場合、16S rRNA検出用に設計されたPCR反应用プライマーを用いてPCR法によって対象核酸を増幅したり、或いはPCR増幅物を元にさらにPCR反応等を行なって調製したりする。また、PCR以外のLAMP法などの増幅方法により調製してもよい。

**【0089】**

その後で、生化学的増幅1202により増幅されたサンプル、またはサンプル1201

そのものに、可視化のために各種標識法により標識する（ラベル混入1203）。この標識物質としては、通常Cy3、Cy5、Rodaminなどの蛍光物質が用いられる。また、生化学的増幅1202において標識分子を混入することもある。

#### 【0090】

こうして標識分子が付加された核酸は、DNAマイクロアレイ1204とハイブリダイゼーション反応（1205）を行う。この様子は、図3に示した通りである。例えば感染症の原因菌の特定をする場合、DNAマイクロアレイ1204は菌に特異的なプローブを基板に固定したものとなる。ここで、各菌に対応したプローブの設計は、例えば16s rRNAをコーディングしているゲノム部分より、上記で説明した方法で行う。なお、DNAマイクロアレイ1204のプローブを固定する担体（基板）は、ガラス基板、プラスチック基板、シリコンウェハー等の平面基板が考えられる。また、凹凸のある三次元構造体、ビーズのような球状のもの、棒状、紐状、糸状のもの等を用いても、本発明の実現形態、効果には影響ない。

#### 【0091】

通常、基板の表面はプローブDNAの固定化が可能なように処理したものが使用される。特に、表面に化学反応が可能となるように官能基を導入した物は、ハイブリダイゼーション反応の過程でプローブが安定に結合している為に、再現性の点で好ましい形態である。本実施形態で用いられる固定化方法は、例えば、マレイミド基とチオール（-SH）基との組み合わせを用いる例が挙げられる。即ち核酸プローブの末端にチオール（-SH）基を結合させておき、固相表面がマレイミド基を有するように処理しておくことで、固相表面に供給された核酸プローブのチオール基と固相表面のマレイミド基とが反応して核酸プローブを固定化する。マレイミド基の導入方法としては、まず、ガラス基板にアミノシランカップリング剤を反応させ、次にそのアミノ基とEMCS試薬（N-(6-Maleimidocaproyloxy)succinimide : Dojin社製）との反応によりマレイミド基を導入する。DNAへのSH基の導入は、DNA自動合成機でDNAを合成する際に5'-Thiol-ModifierC6（Glen Research社製）を用いることにより行なうことができる。固定化に利用する官能基の組み合わせとしては、上記したチオール基とマレイミド基の組み合わせ以外にも、例えばエポキシ基（固相上）とアミノ基（核酸プローブ末端）の組み合わせ等が挙げられる。また、各種シランカップリング剤による表面処理も有効であり、該シランカップリング剤により導入された官能基と反応可能な官能基を導入したオリゴヌクレオチドが用いられる。さらに、官能基を有する樹脂をコーティングする方法も利用可能である。

#### 【0092】

ハイブリダイゼーション反応1205を行った後、DNAマイクロアレイ1204の表面を洗浄し、プローブと結合していない核酸を剥がした後で、（通常は）乾燥し、1205の蛍光量を測定する。ここでは、DNAマイクロアレイ1204の基板に励起光を照射し、蛍光強度を測定した画像を得る（1206、1207）。

#### 【0093】

以下、図19を用いて説明した実験の流れを感染症の原因菌特定の目的を想定した具体的実験操作として説明する。なお、本発明にかかわる生物種類判定方法は、以下に述べる感染症の原因菌特定に限ったものではなく、MHCなどの人間の体質判定や、癌などの疾病に関わるDNA、RNAの解析に用いてもよい。

#### 【0094】

##### <1. プローブDNAの準備>

Enterobacter cloacae株検出用Probeとして配列番号59～65に示す核酸配列（In）（nは数字）を設計した。具体的には、16s rRNAをコーディングしているゲノム部分より、上記第1実施形態或いは第2実施形態で説明したい方法を用いて設計した。なお、データベースは、NCBIデータベースを利用した。

#### 【0095】

配列番号59～65（相補鎖配列は配列番号137～143）に示したプローブは、DNAマイクロアレイに固定するための官能基として、合成後、定法に従って核酸の5'末端

にチオール基を導入した。官能基の導入後、精製し、凍結乾燥した。凍結乾燥したプローブは、 $-30^{\circ}\text{C}$ の冷凍庫に保存した。

#### 【0096】

黄色ブドウ球菌、表皮ブドウ球菌、大腸菌、肺炎桿菌、緑膿菌、セラチア菌、肺炎連鎖球菌、インフルエンザ菌、及びエンテロコッカス・フェカリス菌についても同様な手法により以下に示すプローブセットを設計した。

#### 【0097】

黄色ブドウ球菌：配列番号1～9（相補鎖配列は配列番号79～87）、  
表皮ブドウ球菌：配列番号10～16（相補鎖配列は配列番号88～94）、  
大腸菌：配列番号17～23（相補鎖配列は配列番号95～101）、  
肺炎桿菌：配列番号24～29（相補鎖配列は配列番号102～107）、  
緑膿菌：配列番号30～37（相補鎖配列は配列番号108～115）、  
セラチア菌：配列番号38～43（相補鎖配列は配列番号116～121）、  
肺炎連鎖球菌：配列番号44～50（相補鎖配列は配列番号122～128）、  
インフルエンザ菌：配列番号51～58（相補鎖配列は配列番号129～136）、  
エンテロコッカス・フェカリス菌：配列番号66～72（相補鎖配列は配列番号144～150）。

#### 【0098】

<2. 検体増幅用PCR Primer の準備>

起炎菌検出用の為の16s rRNA核酸（標的核酸）増幅用PCR Primerとして表2に示す核酸配列を設計した。

#### 【0099】

具体的には、16s rRNAをコーディングしているゲノム部分を特異的に増幅するプローブセット、つまり約1500塩基長の16s rRNAコーディング領域の両端部分で、特異的な融解温度をできるだけ揃えたプライマーを設計した。なお、変異株や、ゲノム上に複数存在する16s rRNAコーディング領域も同時に増幅できるように複数種類のプライマーを設計した。

#### 【0100】

##### 【表1】

	Primer No.	配列
Forward Primer	F-1	5' GCGGCGTGCCTAATACATGCAAG 3'
	F-2	5' GCGGCAGGCCTAACACATGCAAG 3'
	F-3	5' GCGGCAGGCTTAACACATGCAAG 3'
Reverse Primer	R-1	5' ATCCAGCCGCACCTTCCGATAC 3'
	R-2	5' ATCCAACCGCAGGTTCCCTAC 3'
	R-3	5' ATCCAGCCGCAGGTTCCCTAC 3'

#### 【0101】

表中に示したPrimerは、合成後、高速液体クロマトグラフィー（HPLC）により精製し、Forward Primerを3種、Reverse Primerを3種混合し、それぞれのPrimer濃度が、最終濃度10 pmol/ $\mu\text{l}$  となるようにTE緩衝液に溶解した。

#### 【0102】

<3. Enterobacter cloacae Genome DNA（モデル検体）の抽出>

<3-1. 微生物の培養 & Genome DNA 抽出の前処理>

まず、Enterobacter cloacae 標準株を、定法に従って培養した。この微生物培養液を1

.5ml容量のマイクロチューブに1.0ml ( $OD_{600}=0.7$ ) 採取し、遠心分離で菌体を回収した (8500rpm、5min、4℃)。上精を捨てた後、Enzyme Buffer (50mM Tris-HCl: p.H. 8.0、25 mM EDTA) 300  $\mu$ lを加え、ミキサーを用いて再懸濁した。再懸濁した菌液は、再度、遠心分離で菌体を回収した (8500rpm、5min、4℃)。上精を捨てた後、回収された菌体に、以下の酵素溶液を加え、ミキサーを用いて再懸濁した。

Lysozyme 50  $\mu$ l (20 mg/ml in Enzyme Buffer)  
N-Acetylmuramidase SG 50  $\mu$ l (0.2 mg/ml in Enzyme Buffer)。

#### 【0103】

次に、酵素溶液を加え再懸濁した菌液を、37℃のインキュベーター内で30分間静置し、細胞壁の溶解処理を行った。

#### 【0104】

##### <3-2. Genome抽出>

以下に示す微生物のGenome DNA抽出は、核酸精製キット (MagExtractor -Genome- : TOYOBO社製) を用いて行った。

#### 【0105】

具体的には、まず、前処理した微生物懸濁液に溶解・吸着液750  $\mu$ lと磁性ビーズ40  $\mu$ lを加え、チューブミキサーを用いて、10分間激しく攪拌した (ステップ1)。

次に、分離用スタンド (Magical Trapper) にマイクロチューブをセットし、30秒間静置して磁性粒子をチューブの壁面に集め、スタンドにセットした状態のまま、上精を捨てた (ステップ2)。

次に、洗浄液 900  $\mu$ l を加え、ミキサーで5sec程度攪拌して再懸濁を行った (ステップ3)。

次に、分離用スタンド (Magical Trapper) にマイクロチューブをセットし、30秒間静置して磁性粒子をチューブの壁面に集め、スタンドにセットした状態のまま、上精を捨てた (ステップ4)。

ステップ3、4を繰り返して2度目の洗浄 (ステップ5) を行った後、70%エタノール 900  $\mu$ l を加え、ミキサーで5sec程度攪拌して再懸濁した (ステップ6)。

#### 【0106】

次に、分離用スタンド (Magical Trapper) にマイクロチューブをセットし、30秒間静置して磁性粒子をチューブの壁面に集め、スタンドにセットした状態のまま、上精を捨てた (ステップ7)。

ステップ6、7を繰り返して70%エタノールによる2度目の洗浄 (ステップ8) を行った後、回収された磁性粒子に純水 100  $\mu$ l を加え、チューブミキサーで10分間攪拌を行った。

次に分離用スタンド (Magical Trapper) にマイクロチューブをセットし、30秒間静置して磁性粒子をチューブ壁面に集め、スタンドにセットした状態のまま、上精を新しいチューブに回収した。

#### 【0107】

##### <3-3. 回収したGenome DNAの検査>

回収された微生物 (*Enterobacter cloacae* 株) のGenome DNAは、定法に従って、アガロース電気泳動と260/280nmの吸光度測定を行い、その品質 (低分子核酸の混入量、分解の程度) と回収量を検定した。

#### 【0108】

本実験操作では、約10  $\mu$ gのGenome DNA が回収され、Genome DNAのデグラデーションやrRNAの混入は認められなかった。回収したGenome DNAは、最終濃度50ng/ $\mu$ lとなるようにTE緩衝液に溶解し、以下の実験操作に使用した。

#### 【0109】

##### <4. DNAマイクロアレイの作製>

##### <4-1. ガラス基板の洗浄>

合成石英のガラス基板 (サイズ: 25mm x 75mm x 1mm、飯山特殊ガラス社製) を耐熱、耐



アルカリ のラックに入れ、所定の濃度に調製した超音波洗浄用の洗浄液に浸した。一晚洗浄液中で浸した後、20分間超音波洗浄を行った。続いて基板を取り出し、軽く純水ですすいだ後、超純水中で20分超音波洗浄をおこなった。次に80℃に加熱した1N水酸化ナトリウム水溶液中に10分間基板を浸した。再び純水洗浄と超純水洗浄を行い、DNAチップ用の石英ガラス基板を用意した。

#### 【0110】

##### <4-2. 表面処理>

シランカップリング剤KBM-603(信越シリコン社製)を、1%の濃度となるように純水中に溶解させ、2時間室温で攪拌した。続いて、先に洗浄したガラス基板をシランカップリング剤水溶液に浸し、20分間室温で放置した。ガラス基板を引き上げ、軽く純水で表面を洗浄した後、窒素ガスを基板の両面に吹き付けて乾燥させた。次に乾燥した基板を120℃に加熱したオーブン中で1時間ベークし、カップリング剤処理を完結させ、基板表面にアミノ基を導入した。次いで同仁化学研究所社製のN-マレイミドカプロイロキシスクシイミド(N-(6-Maleimidocaproyloxy)succinimido) (以下EMCSと略す)を、ジメチルスルホキシドとエタノールの1:1混合溶媒中に最終濃度が0.3mg/mlとなるように溶解したEMCS溶液を用意した。ベークの終了したガラス基板を放冷し、調製したEMCS溶液中に室温で2時間浸した。この処理により、シランカップリング剤によって表面に導入されたアミノ基とEMCSのスクシイミド基が反応し、ガラス基板表面にマレイミド基が導入された。EMCS溶液から引き上げたガラス基板を、先述のMCSを溶解した混合溶媒を用いて洗浄し、さらにエタノールにより洗浄した後、窒素ガス雰囲気下で乾燥させた。

#### 【0111】

##### <4-3. プローブDNA>

実験操作1で作製した微生物検出用プローブを純水に溶解し、それぞれ、最終濃度(インク溶解時)10 $\mu$ Mとなるように分注した後、凍結乾燥を行い、水分を除いた。

#### 【0112】

##### <4-4. BJプリンターによるDNA吐出、および基板への結合>

グリセリン7.5wt%、チオジグリコール7.5wt%、尿素7.5wt%、アセチレノールEH(川研ファインケミカル社製)1.0wt%を含む水溶液を用意した。続いて、先に用意した7種類のプローブ(表1)を上記の混合溶媒に規定濃度となるように溶解した。得られたDNA溶液をバブルジェット(登録商標)プリンター(商品名:BJF-850 キヤノン社製)用インクタンクに充填し、印字ヘッドに装着した。

#### 【0113】

なおここで用いたバブルジェット(登録商標)プリンターは平板への印刷が可能のように改造を施したものである。またこのバブルジェット(登録商標)プリンターは、所定のファイル作成方法に従って印字パターンを入力することにより、約5ピコリットルのDNA溶液を約120マイクロメートルピッチでスポッティングすることが可能となっている。

#### 【0114】

続いて、この改造バブルジェット(登録商標)プリンターを用いて、1枚のガラス基板に対して、印字操作を行い、アレイを作製した。印字が確実に行われていることを確認した後、30分間加湿チャンバー内に静置し、ガラス基板表面のマレイミド基と核酸プローブ末端のチオール基とを反応させた。

#### 【0115】

##### <4-5. 洗浄>

30分間の反応後、100mMのNaClを含む10mMのリン酸緩衝液(pH7.0)により表面に残ったDNA溶液を洗い流し、ガラス基板表面に一本鎖DNAが固定したDNAマイクロアレイを得た。

#### 【0116】

##### <5. 検体の増幅と標識化(PCR増幅&蛍光標識の取り込み)>

検体となる微生物DNAの増幅、および、標識化反応を以下に示す。

#### 【0117】



【表 2】

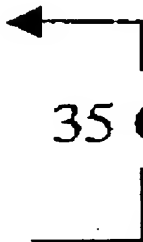
Premix PCR 試薬 (TAKARA ExTaq)	25 $\mu$ l	
Template Genome DNA	2 $\mu$ l	(100ng)
Forward Primer mix	2 $\mu$ l	(20pmol/tube each)
Reverse Primer mix	2 $\mu$ l	(20pmol/tube each)
Cy-3 dUTP (1mM)	2 $\mu$ l	(2nmol/tube)
H <sub>2</sub> O	17 $\mu$ l	
Total	50 $\mu$ l	

【0118】

上記組成の反応液を以下のプロトコールに従って、市販のサーマルサイクラーで増幅反応を行った。

【0119】

【表 3】

95°C	10 min.	
92°C	45 sec.	 35 Cycles
55°C	45 sec.	
72°C	45 sec.	
72°C	10 min.	

【0120】

反応終了後、精製用カラム (QIAGEN QIAquick PCR Purification Kit) を用いて Primer を除去した後、増幅産物の定量を行い、標識化検体とした。

【0121】

&lt; 6. ハイブリダイゼーション &gt;

< 4. DNA マイクロアレイの作製 > で作製した DNA マイクロアレイと < 5. 検体の増幅と標識化 (PCR 増幅 & 蛍光標識の取り込み) > で作製した標識化検体を用いて検出反応を行った。

【0122】

## &lt;6-1. DNAマイクロアレイのブロッキング&gt;

BSA (牛血清アルブミンFraction V: Sigma社製) を1wt%となるように100mM NaCl / 10mM Phosphate Bufferに溶解し、この溶液にDNAマイクロアレイの作製で作製したDNAマイクロアレイを室温で2時間浸し、ブロッキングを行った。ブロッキング終了後、0.1wt% SDS (ドデシル硫酸ナトリウム) を含む2x SSC溶液 (NaCl 300mM、Sodium Citrate (trisodium citrate dihydrate,  $C_6H_5Na_3 \cdot 2H_2O$ ) 30mM、p.H. 7.0) で洗浄を行った後、純水でリンスしてからスピンドライ装置で水切りを行った。

## 【0123】

## &lt;6-2. ハイブリダイゼーション&gt;

水切りしたDNAマイクロアレイをハイブリダイゼーション装置 (Genomic Solutions Inc. Hybridization Station) にセットし、以下に示すハイブリダイゼーション溶液、条件でハイブリダイゼーション反応を行った。

## 【0124】

## ・ハイブリダイゼーション溶液

6x SSPE / 10% Form amide / Target (2nd PCR Products 全量)  
(6xSSPE: NaCl 900mM、 $NaH_2PO_4 \cdot H_2O$  60mM、EDTA 6mM、p.H. 7.4)

## ・ハイブリダイゼーション条件

65℃ 3min → 92℃ 2min → 45℃ 3hr → Wash 2x SSC / 0.1% SDS at 25℃ → Wash 2x SSC at 20℃ → (Rinse with  $H_2O$  : Manual) → Spin dry (65℃で3分、92度で2分、45℃で3時間ハイブリダイゼーション反応させた後、2x SSC / 0.1% SDS、25℃で洗浄、2xSSC、20℃で洗浄後、純水でリンスしスピンドライした)。

## 【0125】

## &lt;7. 微生物の検出 (蛍光測定) &gt;

ハイブリダイゼーション反応終了後のDNAマイクロアレイをDNAマイクロアレイ用蛍光検出装置 (Axon社製、GenePix 4000B) を用いて蛍光測定を行った。夫々のプローブで良好な判別結果が得られた。

## 【0126】

以上説明したように、本実施形態によれば、DNAマイクロアレイシステムに最適なオリゴヌクレオチドプローブの設計が実現できる。このことにより、より確かな生物種、個体識別情報が得られるという効果がある。

## 【0127】

## &lt;他の実施形態&gt;

なお、本発明の目的は、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記憶媒体を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ (またはCPUやMPU) が記憶媒体に格納されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。

## 【0128】

この場合、記憶媒体から読出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。

## 【0129】

プログラムコードを供給するための記憶媒体としては、例えば、フレキシブルディスク、ハードディスク、光ディスク、光磁気ディスク、CD-ROM、CD-R、磁気テープ、不揮発性のメモリカード、ROMなどを用いることができる。

## 【0130】

また、コンピュータが読出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているOS (オペレーティングシステム) などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

## 【0131】

さらに、記憶媒体から読出されたプログラムコードが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込まれた後、そのプログラムコードの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

## 【図面の簡単な説明】

## 【0132】

【図1】第1実施形態によるプローブ設計方法の概要を説明する図である。

【図2】第1実施形態によるプローブ設計方法が適用可能な情報処理装置の構成を示すブロック図である。

【図3】ハイブリダイゼーション反応を説明する図である。

【図4】感染症の判定用DNAマイクロアレイの実験手順を説明する図である。

【図5】黄色ブドウ球菌の異なる株のゲノム構造を説明する図である。

【図6】実施形態による頻度テーブルの例を示す図である。

【図7】ターゲット塩基配列の特異性をスキャンした例を示す図である。

【図8】競合頻度テーブルと自己頻度テーブルの値をグラフ化した図である。

【図9】第1実施形態によるプローブ設計方法を説明するフローチャートである。

【図10】第1実施形態によるプローブ設計方法におけるユーザインターフェースを示す図である。

【図11】第1実施形態によるプローブ設計方法におけるユーザインターフェースを示す図である。

【図12】第2実施形態によるプローブ設計方法の概要を説明する図である。

【図13】ターゲット塩基配列に存在する共通領域と、その共通領域に挟まれた部分を表す図である。

【図14】第2実施形態によるプローブ設計方法を説明するフローチャートである。

【図15】第2実施形態によるプローブ設計方法におけるユーザインターフェースを示す図である。

【図16】第2実施形態によるプローブ設計方法におけるユーザインターフェースを示す図である。

【図17】第2実施形態によるプローブの自動設計方法を説明する図である。

【図18】第2実施形態によるプローブの自動設計方法を説明するフローチャートである。

【図19】実施形態のプローブ設計方法で設計したプローブを用いて核酸解析を行う手順を説明する図である。

## 【配列表】

## SEQUENCE LISTING

<110> CANON INC.

<120> DNA PROBE DESIGNING APPARATUS AND INFORMATION PROCESSING METHOD  
THEREOF

<130> 5516222-01

<150> JPA2003-099464

<151> 2003-04-02

<160> 170

<170> PatentIn version 3.2

<210> 1

<211> 26

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 1

gaaccgcatg gttcaaaagt gaaaga

26

<210> 2

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 2

cacttataga tggatccgcg ctgc

24

<210> 3

<211> 26

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 3

tgcacatctt gacggtacct aatcag

26

<210> 4

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 4

ccccttagtg ctgcagctaa cg

22

<210> 5

<211> 23

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 5

aatacaaagg gcagcgaaac cgc

23

<210> 6

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 6

ccggtggagt aaccttttag gagct

25

<210> 7

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 7

taacctttta ggagctagcc gtcga

25

<210> 8  
<211> 23  
<212> DNA  
<213> Artificial  
  
<220>  
<223> Synthesized DNA probe  
  
<400> 8  
tttaggagct agccgtcgaa ggt 23

<210> 9  
<211> 23  
<212> DNA  
<213> Artificial  
  
<220>  
<223> Synthesized DNA probe  
  
<400> 9  
tagccgtcga aggtgggaca aat 23

<210> 10  
<211> 22  
<212> DNA  
<213> Artificial  
  
<220>  
<223> Synthesized DNA probe  
  
<400> 10  
gaacagacga ggagcttgct cc 22

<210> 11  
<211> 26  
<212> DNA  
<213> Artificial  
  
<220>  
<223> Synthesized DNA probe  
  
<400> 11  
tagtgaaaga cggttttgct gtcact 26

<210> 12

<211> 26  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 12  
taagtaacta tgcacgtctt gacggt 26

<210> 13  
<211> 26  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 13  
gaccctcta gagatagagt tttccc 26

<210> 14  
<211> 24  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 14  
agtaaccatt tggagctagc cgtc 24

<210> 15  
<211> 23  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 15  
gagcttgctc ctctgacgtt agc 23

<210> 16  
<211> 22  
<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 16

agccggtgga gtaaccattt gg

22

<210> 17

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 17

ctcttgccat cggatgtgcc ca

22

<210> 18

<211> 26

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 18

atacctttgc tcattgacgt tacccg

26

<210> 19

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 19

tttgctcatt gacgttaccc gcag

24

<210> 20

<211> 24

<212> DNA

<213> Artificial



<220>  
<223> Synthesized DNA probe  
  
<400> 20  
actggcaagc ttgagtctcg taga 24

<210> 21  
<211> 23  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe  
  
<400> 21  
atacaaagag aagcgacctc gcg 23

<210> 22  
<211> 25  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe  
  
<400> 22  
cggacctcat aaagtgcgtc gtagt 25

<210> 23  
<211> 26  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe  
  
<400> 23  
gcggggagga agggagtaaa gttaat 26

<210> 24  
<211> 23  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 24  
tagcacagag agcttgctct cgg 23

<210> 25  
<211> 23  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 25  
tcatgccatc agatgtgccc aga 23

<210> 26  
<211> 25  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 26  
cggggaggaa ggcgataagg ttaat 25

<210> 27  
<211> 25  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 27  
ttcgattgac gttacccgca gaaga 25

<210> 28  
<211> 26  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 28

ggtctgtcaa gtcggatgtg aaatcc

26

<210> 29

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 29

gcaggctaga gtcttgtaga gggg

24

<210> 30

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 30

tgagggagaa agtgggggat cttc

24

<210> 31

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 31

tcagatgagc ctaggtcgga ttagc

25

<210> 32

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 32

gagctagagt acggtagagg gtgg

24

<210> 33  
<211> 25  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 33  
gtacggtaga ggggtggtgga atttc 25

<210> 34  
<211> 24  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 34  
gaccacctgg actgatactg acac 24

<210> 35  
<211> 25  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 35  
tggccttgac atgctgagaa ctttc 25

<210> 36  
<211> 23  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 36  
ttagttacca gcacctcggg tgg 23

<210> 37

<211> 22  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 37  
tagtctaacc gcaaggggga cg 22

<210> 38  
<211> 23  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 38  
tagcacaggg gagcttgctc cct 23

<210> 39  
<211> 25  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 39  
aggtggtgaa ctaatacgc tcac 25

<210> 40  
<211> 26  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 40  
tcacaaattg acgttactcg cagaag 26

<210> 41  
<211> 26  
<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 41

actgcatttg aaactggcaa gctaga

26

<210> 42

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 42

ttatcctttg ttgccagcgg tt

22

<210> 43

<211> 23

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 43

actttcagcg aggaggaagg tgg

23

<210> 44

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 44

agtagaacgc tgaaggagga gcttg

25

<210> 45

<211> 25

<212> DNA

<213> Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 45

cttgcacac taccagatgg acctg

25

&lt;210&gt; 46

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 46

tgagagtgga aagttcacac tgtgac

26

&lt;210&gt; 47

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 47

gctgtggcctt aaccatagta ggcttt

26

&lt;210&gt; 48

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 48

aagcggctct ctggcttgta act

23

&lt;210&gt; 49

&lt;211&gt; 24

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

<400> 49  
tagacccttt ccggggttta gtgc 24

<210> 50  
<211> 26  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 50  
gacggcaagc taatctctta aagcca 26

<210> 51  
<211> 24  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 51  
gcttggaat ctggcttatg gagg 24

<210> 52  
<211> 23  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 52  
tgccatagga tgagccaag tgg 23

<210> 53  
<211> 26  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 53



● cttgggaatg tactgacgct catgtg

26

&lt;210&gt; 54

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 54

ggattgggct tagagcttgg tgc

23

&lt;210&gt; 55

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 55

tacagaggga agcgaagctg cg

22

&lt;210&gt; 56

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 56

ggcgtttacc acggtatgat tcatga

26

&lt;210&gt; 57

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 57

aatgcctacc aagcctgcga tct

23

<210> 58  
<211> 25  
<212> DNA  
<213> Artificial  
  
<220>  
<223> Synthesized DNA probe  
  
<400> 58  
tatcggaaga tgaaagtgcg ggact 25  
  
<210> 59  
<211> 22  
<212> DNA  
<213> Artificial  
  
<220>  
<223> Synthesized DNA probe  
  
<400> 59  
cagagagctt gctctcgggt ga 22  
  
<210> 60  
<211> 26  
<212> DNA  
<213> Artificial  
  
<220>  
<223> Synthesized DNA probe  
  
<400> 60  
gggaggaagg tggtgtggtt aataac 26  
  
<210> 61  
<211> 26  
<212> DNA  
<213> Artificial  
  
<220>  
<223> Synthesized DNA probe  
  
<400> 61  
ggtgttgtgg ttaataacca cagcaa 26  
  
<210> 62

<211> 22  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 62  
gcggtctgtc aagtcggatg tg 22

<210> 63  
<211> 25  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 63  
attcgaaact ggcaggctag agtct 25

<210> 64  
<211> 25  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 64  
taaccacagc aattgacgtt acccg 25

<210> 65  
<211> 24  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 65  
gcaattgacg ttacccgcag aaga 24

<210> 66  
<211> 23  
<212> DNA

● <213> Artificial

<220>

<223> Synthesized DNA probe

<400> 66

ttctttcctc ccgagtgcctt gca

23

<210> 67

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 67

aacacgtggg taacctaccc atcag

25

<210> 68

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 68

atggcataag agtgaaaggc gctt

24

<210> 69

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 69

gacccgcggt gcattagcta gt

22

<210> 70

<211> 25

<212> DNA

<213> Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 70

ggacgttagt aactgaacgt cccct

25

&lt;210&gt; 71

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 71

ctcaaccggg gagggtcatt gg

22

&lt;210&gt; 72

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 72

ttggagggtt tccgcccttc ag

22

&lt;210&gt; 73

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA for forward primer named F1

&lt;400&gt; 73

gcggcgtgcc taatacatgc aag

23

&lt;210&gt; 74

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA for forward primer named F2

<400> 74  
gcggcaggcc taacacatgc aag 23

<210> 75  
<211> 23  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA for forward primer named F3

<400> 75  
gcggcaggct taacacatgc aag 23

<210> 76  
<211> 22  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA for reverse primer named R1

<400> 76  
atccagccgc accttccgat ac 22

<210> 77  
<211> 22  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA for reverse primer named R2

<400> 77  
atccaaccgc aggttcccct ac 22

<210> 78  
<211> 22  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA for reverse primer named R3

<400> 78

atccagccgc aggttcccct ac

22

<210> 79

<211> 26

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 79

tctttcactt ttgaaccatg cggttc

26

<210> 80

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 80

gcagcgcgga tccatctata agtg

24

<210> 81

<211> 26

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 81

ctgattaggt accgtcaaga tgtgca

26

<210> 82

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 82

cgtttagctgc agcactaagg gg

22

<210> 83  
<211> 23  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 83  
gcggttttcgc tgccctttgt att 23

<210> 84  
<211> 25  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 84  
agctcctaaa aggttactcc accgg 25

<210> 85  
<211> 25  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 85  
tcgacggcta gctcctaaaa ggtta 25

<210> 86  
<211> 23  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 86  
accttcgacg gctagctcct aaa 23

<210> 87



<211> 23  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 87  
atttgtccca ccttcgacgg cta

23

<210> 88  
<211> 22  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 88  
ggagcaagct cctcgtctgt tc

22

<210> 89  
<211> 26  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 89  
agtgacagca aaaccgtctt tcacta

26

<210> 90  
<211> 26  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 90  
accgtcaaga cgtgcatagt tactta

26

<210> 91  
<211> 26  
<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 91

gggaaaactc tatctctaga ggggtc

26

<210> 92

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 92

gacggctagc tccaaatggt tact

24

<210> 93

<211> 23

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 93

gctaacgtca gaggagcaag ctc

23

<210> 94

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 94

ccaaatggtt actccaccgg ct

22

<210> 95

<211> 22

<212> DNA

<213> Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 95

tgggcacatc cgatggcaag ag

22

&lt;210&gt; 96

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 96

cgggtaacgt caatgagcaa aggtat

26

&lt;210&gt; 97

&lt;211&gt; 24

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 97

ctgcgggtaa cgtcaatgag caaa

24

&lt;210&gt; 98

&lt;211&gt; 24

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 98

tctacgagac tcaagcttgc cagt

24

&lt;210&gt; 99

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

<400> 99  
cgcgaggtcg cttctctttg tat 23

<210> 100  
<211> 25  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 100  
actacgacgc actttatgag gtccg 25

<210> 101  
<211> 26  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 101  
attaacttta ctcccttcct cccgc 26

<210> 102  
<211> 23  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 102  
ccgagagcaa gctctctgtg cta 23

<210> 103  
<211> 23  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 103

tctgggcaca tctgatggca tga

23

<210> 104

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 104

attaacctta tcgccttcct ccccg

25

<210> 105

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 105

tcttctgcgg gtaacgtcaa tcgaa

25

<210> 106

<211> 26

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 106

ggatttcaca tccgacttga cagacc

26

<210> 107

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 107

cccctctaca agactctagc ctgc

24

<210> 108

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 108

gaagatcccc cactttctcc ctca

24

<210> 109

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 109

gctaattccga cctaggctca tctga

25

<210> 110

<211> 24

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 110

ccaccttcta ccgtactcta gctc

24

<210> 111

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 111

gaaattccac caccctctac cgtac

25

<210> 112

<211> 24  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 112  
gtgtcagtat cagtccaggt ggtc

24

<210> 113  
<211> 25  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 113  
gaaagttctc agcatgtcaa ggcca

25

<210> 114  
<211> 23  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 114  
ccacccgagg tgctggtaac taa

23

<210> 115  
<211> 22  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 115  
cgtccccctt gcggttagac ta

22

<210> 116  
<211> 23  
<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 116

agggagcaag ctcccctgtg cta

23

<210> 117

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 117

gatgagcgta ttaagttcac cacct

25

<210> 118

<211> 26

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 118

cttctgcgag taacgtcaat tgatga

26

<210> 119

<211> 26

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 119

tctagcttgc cagtttcaaa tgcagt

26

<210> 120

<211> 22

<212> DNA

<213> Artificial



&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 120

aaccgctggc aacaaaggat aa

22

&lt;210&gt; 121

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 121

ccaccttcct cctcgctgaa agt

23

&lt;210&gt; 122

&lt;211&gt; 25

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 122

caagctcctc cttcagcggt ctact

25

&lt;210&gt; 123

&lt;211&gt; 25

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 123

caggtccatc tggtagtgat gcaag

25

&lt;210&gt; 124

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

<400> 124  
gtcacagtgt gaactttcca ctctca 26

<210> 125  
<211> 26  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 125  
aaagcctact atggttaagc cacagc 26

<210> 126  
<211> 23  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 126  
agttacaagc cagagagccg ctt 23

<210> 127  
<211> 24  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 127  
gcactaaacc ccggaaggg tcta 24

<210> 128  
<211> 26  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 128

●  
tggctttaag agattagctt gccgtc

26

&lt;210&gt; 129

&lt;211&gt; 24

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 129

cctccataag ccagattccc aagc

24

&lt;210&gt; 130

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 130

ccacttgggc tcatacctatg gca

23

&lt;210&gt; 131

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 131

cacatgagcg tcagtacatt cccaag

26

&lt;210&gt; 132

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 132

gcaccaagct ctaagcccaa tcc

23

<210> 133

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 133

cgcagcttcg cttccctctg ta

22

<210> 134

<211> 26

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 134

tcatgaatca taccgtggta aacgcc

26

<210> 135

<211> 23

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 135

agatcgcagg cttggtaggc att

23

<210> 136

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 136

agtcccgcac tttcatcttc cgata

25

<210> 137

<211> 23  
<212> DNA  
<213> Artificial  
  
<220>  
<223> Synthesized DNA probe  
  
<400> 137  
tgcaagcact cgggaggaaa gaa 23

<210> 138  
<211> 25  
<212> DNA  
<213> Artificial  
  
<220>  
<223> Synthesized DNA probe  
  
<400> 138  
ctgatgggta ggttaccac gtgtt 25

<210> 139  
<211> 24  
<212> DNA  
<213> Artificial  
  
<220>  
<223> Synthesized DNA probe  
  
<400> 139  
aagcgccttt cactcttatg ccat 24

<210> 140  
<211> 22  
<212> DNA  
<213> Artificial  
  
<220>  
<223> Synthesized DNA probe  
  
<400> 140  
actagctaat gcaccgcggg tc 22

<210> 141  
<211> 25  
<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 141

aggggacgtt cagttactaa cgtcc

25

<210> 142

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 142

ccaatgaccc tccccggttg ag

22

<210> 143

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 143

ctgaagggcg gaaaccctcc aa

22

<210> 144

<211> 22

<212> DNA

<213> Artificial

<220>

<223> Synthesized DNA probe

<400> 144

tcacccgaga gcaagctctc tg

22

<210> 145

<211> 26

<212> DNA

<213> Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 145

gttattaacc acaacacctt cctccc

26

&lt;210&gt; 146

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 146

ttgctgtggt tattaaccac aacacc

26

&lt;210&gt; 147

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 147

cacatccgac ttgacagacc gc

22

&lt;210&gt; 148

&lt;211&gt; 25

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

&lt;400&gt; 148

agactctagc ctgccagttt cgaat

25

&lt;210&gt; 149

&lt;211&gt; 25

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; Synthesized DNA probe

<400> 149  
cgggtaacgt caattgctgt ggta 25

<210> 150  
<211> 24  
<212> DNA  
<213> Artificial

<220>  
<223> Synthesized DNA probe

<400> 150  
tcttctgcgg gtaacgtcaa ttgc 24

<210> 151  
<211> 10  
<212> DNA  
<213> Artificial

<220>  
<223> DNA exemplified in the drawing

<400> 151  
tgatgtgaaa 10

<210> 152  
<211> 24  
<212> DNA  
<213> Artificial

<220>  
<223> DNA exemplified in the drawing

<400> 152  
gcccacggct caaccgtgga gggt 24

<210> 153  
<211> 10  
<212> DNA  
<213> Artificial

<220>  
<223> DNA exemplified in the drawing

<400> 153



aaactggaaa

10

&lt;210&gt; 154

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; DNA exemplified in the drawing

&lt;400&gt; 154

gaaccgcatg gttcaaaagt gaaaga

26

&lt;210&gt; 155

&lt;211&gt; 24

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; DNA exemplified in the drawing

&lt;400&gt; 155

cacttataga tggatccgcg ctgc

24

&lt;210&gt; 156

&lt;211&gt; 26

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; DNA exemplified in the drawing

&lt;400&gt; 156

tgcacatctt gacggtacct aatcag

26

&lt;210&gt; 157

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial

&lt;220&gt;

&lt;223&gt; DNA exemplified in the drawing

&lt;400&gt; 157

ccccttagtg ctgcagctaa cg

22

<210> 158  
<211> 23  
<212> DNA  
<213> Artificial  
  
<220>  
<223> DNA exemplified in the drawing  
  
<400> 158  
aatacaaagg gcagcgaaac cgc 23  
  
<210> 159  
<211> 25  
<212> DNA  
<213> Artificial  
  
<220>  
<223> DNA exemplified in the drawing  
  
<400> 159  
ccggtggagt aaccttttag gagct 25  
  
<210> 160  
<211> 25  
<212> DNA  
<213> Artificial  
  
<220>  
<223> DNA exemplified in the drawing  
  
<400> 160  
taacctttta ggagctagcc gtcga 25  
  
<210> 161  
<211> 23  
<212> DNA  
<213> Artificial  
  
<220>  
<223> DNA exemplified in the drawing  
  
<400> 161  
tttaggagct agccgtcgaa ggt 23  
  
<210> 162

<211> 23  
<212> DNA  
<213> Artificial

<220>  
<223> DNA exemplified in the drawing

<400> 162  
tagccgtcga aggtgggaca aat 23

<210> 163  
<211> 10  
<212> DNA  
<213> Artificial

<220>  
<223> DNA exemplified in the drawing

<400> 163  
agtcgagcga 10

<210> 164  
<211> 10  
<212> DNA  
<213> Artificial

<220>  
<223> DNA exemplified in the drawing

<400> 164  
gcggcggacg 10

<210> 165  
<211> 30  
<212> DNA  
<213> Artificial

<220>  
<223> DNA exemplified in the drawing

<400> 165  
acggacgaga agcttgcttc tctgatgta 30

<210> 166  
<211> 23  
<212> DNA

<213> Artificial

<220>

<223> DNA exemplified in the drawing

<400> 166

acggacgaga agcttgcttc tct

23

<210> 167

<211> 25

<212> DNA

<213> Artificial

<220>

<223> DNA exemplified in the drawing

<400> 167

tgtcacttat agatggatcc gcgct

25

<210> 168

<211> 26

<212> DNA

<213> Artificial

<220>

<223> DNA exemplified in the drawing

<400> 168

tgtaagtaac tgtgcacatc ttgacg

26

<210> 169

<211> 26

<212> DNA

<213> Artificial

<220>

<223> DNA exemplified in the drawing

<400> 169

acaactctag agatagagcc ttcccc

26

<210> 170

<211> 26

<212> DNA

<213> Artificial

<220>

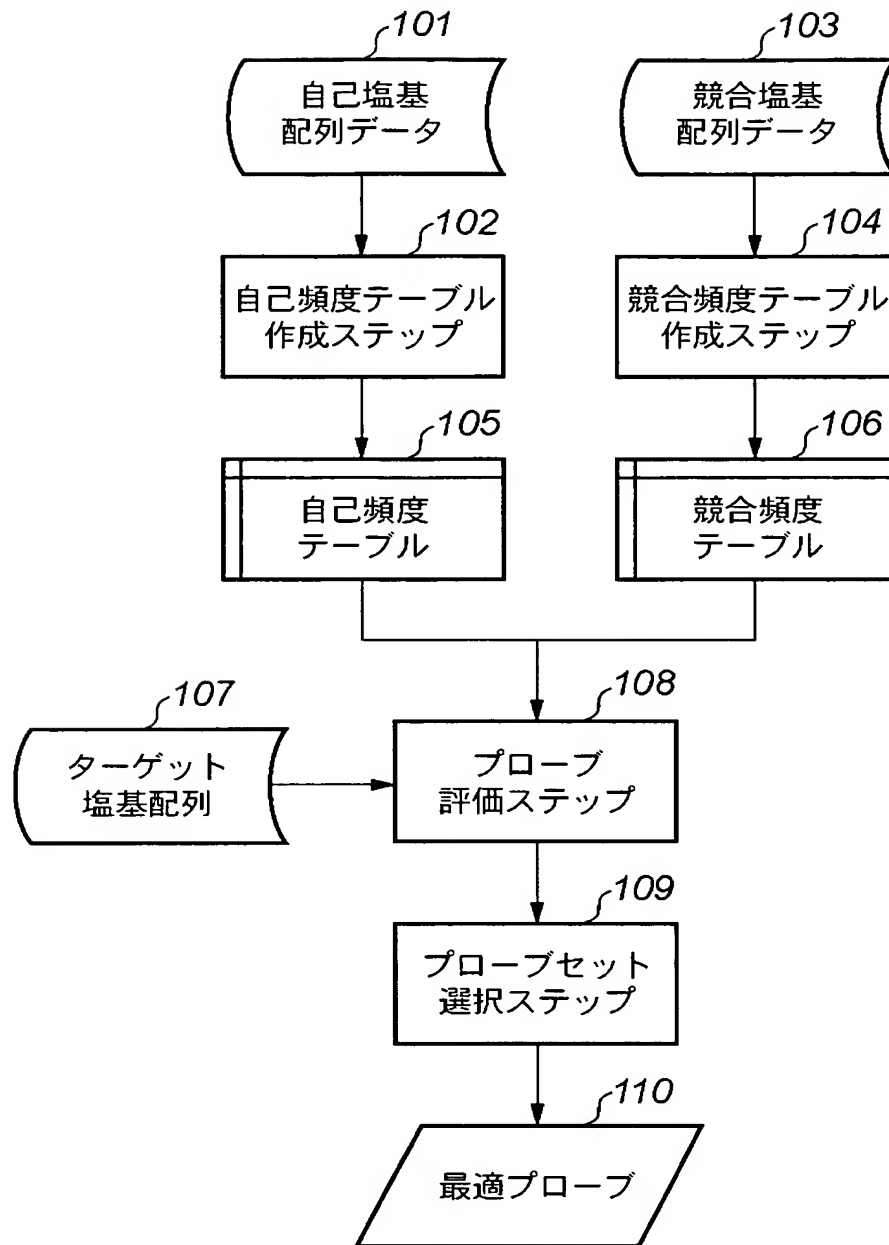
<223> DNA exemplified in the drawing

<400> 170

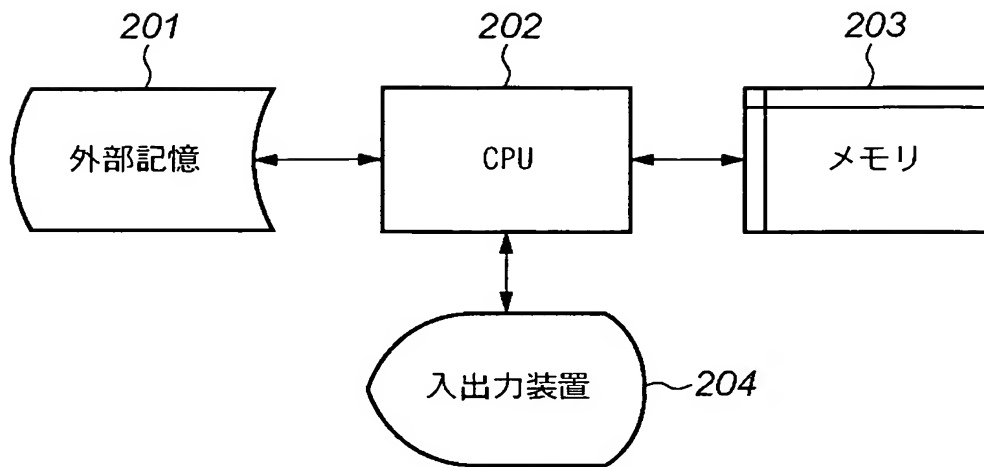
gtggagtaac cttttaggag ctagcc

26

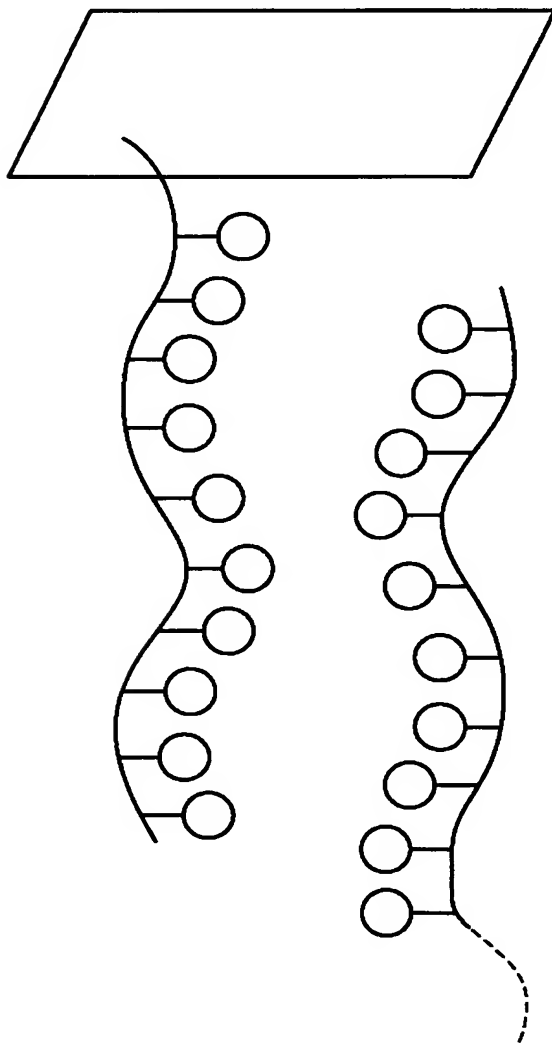
【書類名】図面  
【図1】



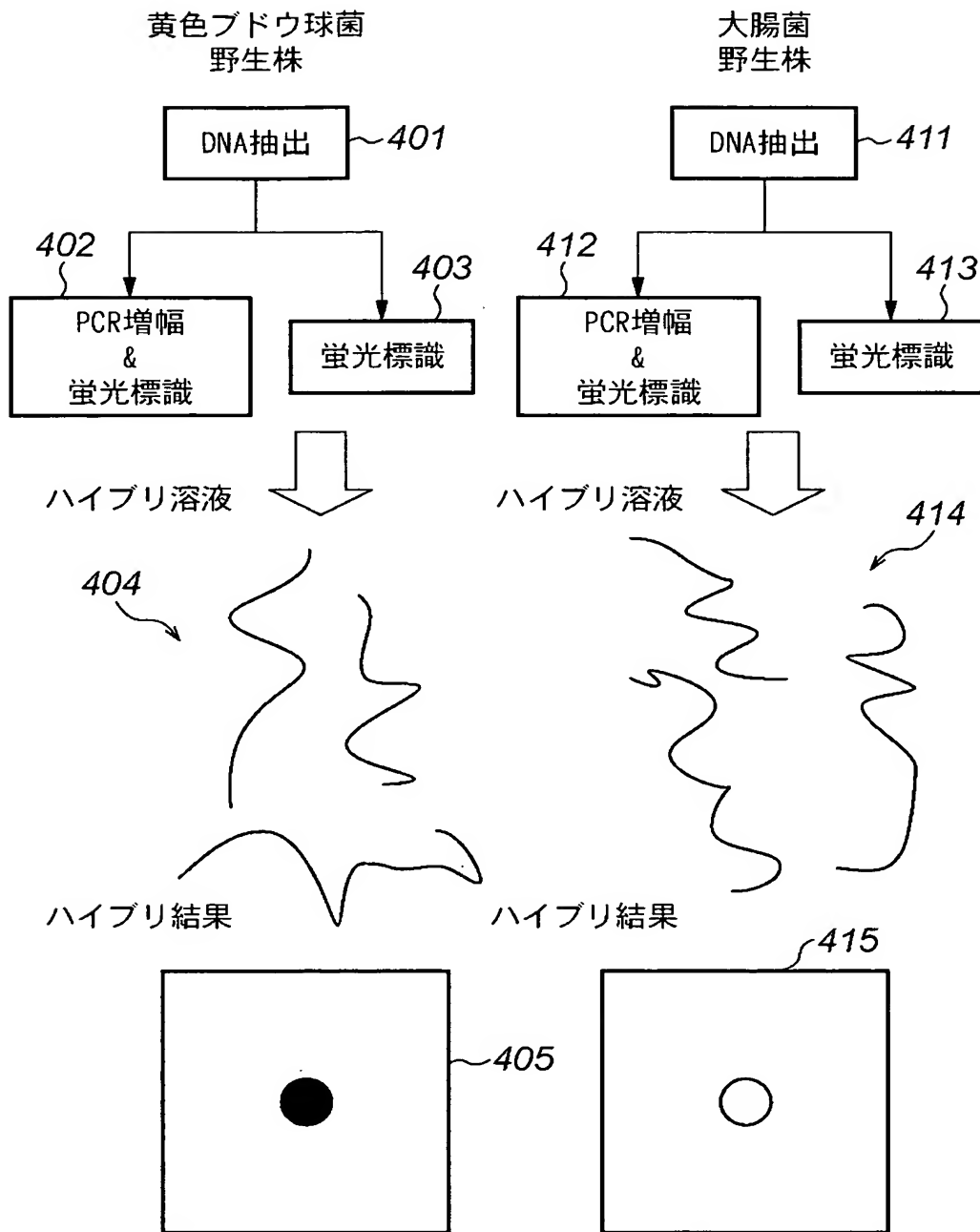
【図 2】



【図 3】

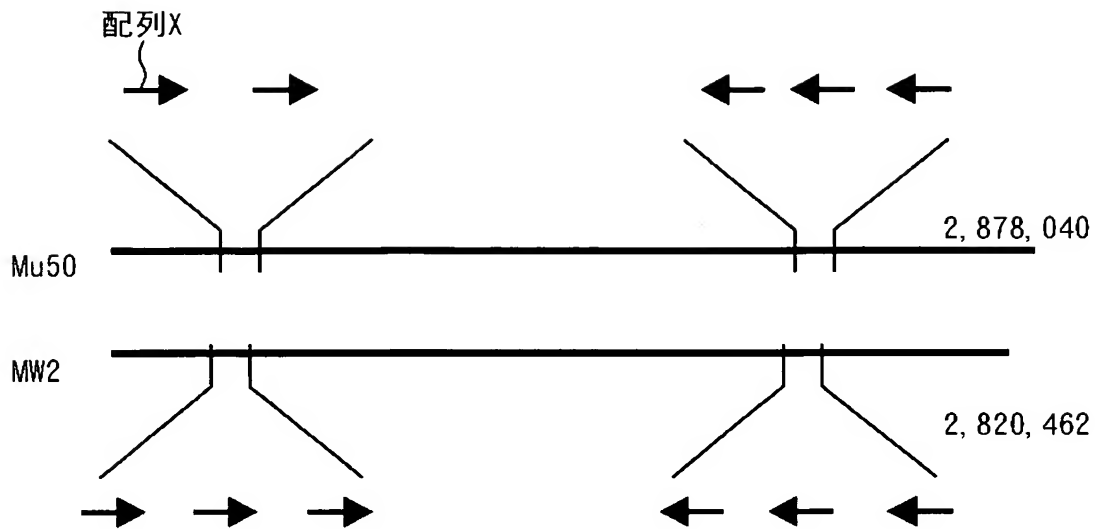


【図 4】





【図 5】



【図 6】

塩基配列	頻度(一特異性)
AAAAAAAAA	1000
AAAAAAAAAT	10
...	...
...	...
GTAGATCCA	3
GTAGATCCT	5
...	...
...	...
CCCCCCCCC	0

【図 7】

ターゲット塩基配列

ATGCGTA.....GTAAA

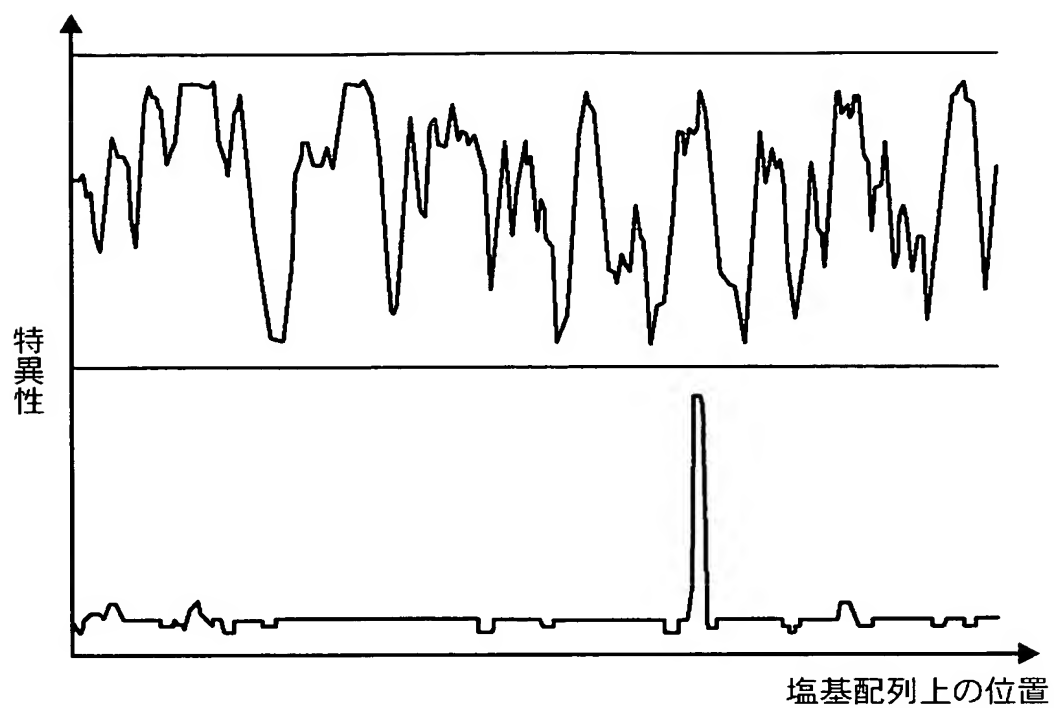
特異性

ATGCGTACG  
TGGGTACGA  
GCGTACGAC  
CGTACGACT

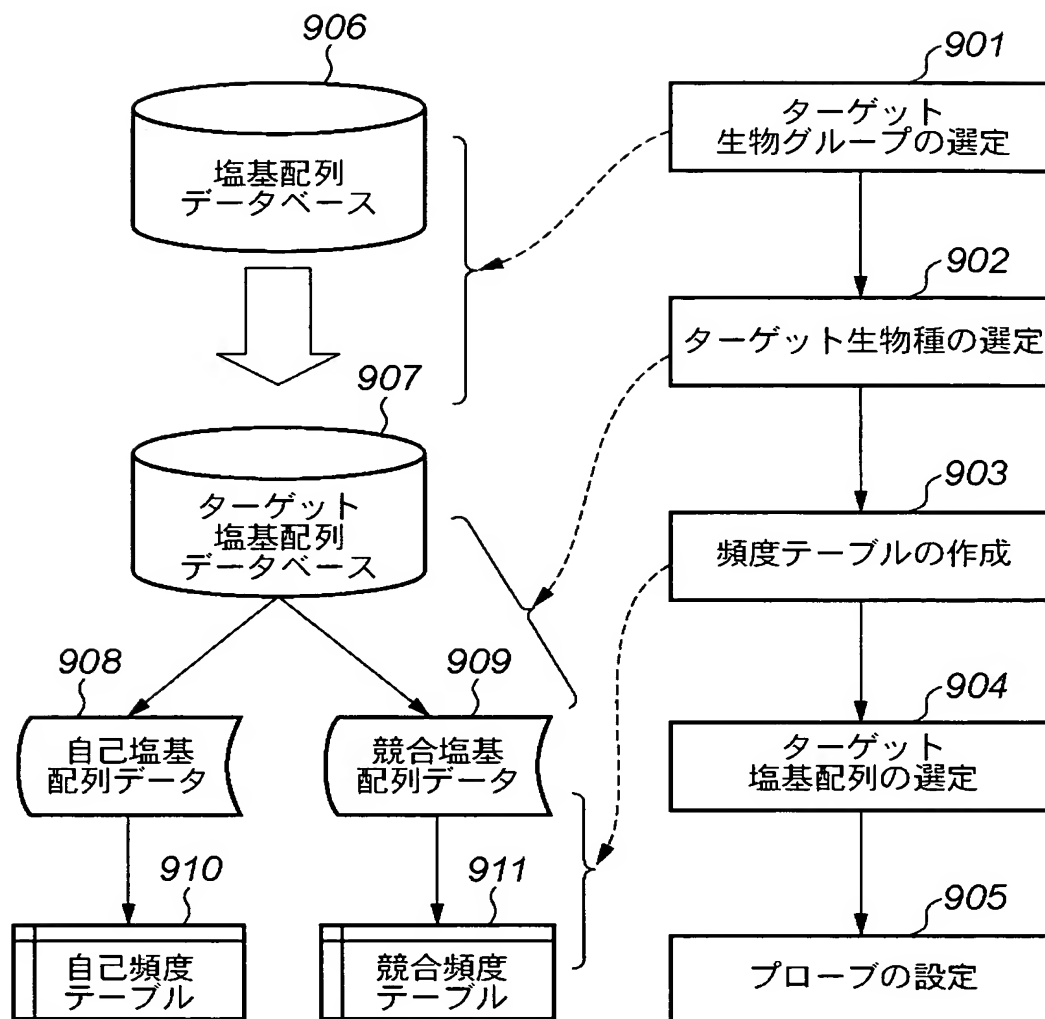
-300  
-250  
-200  
-220  
-400  
...  
...  
-300  
-350

TATGCGTAA  
ATGCGTAAA

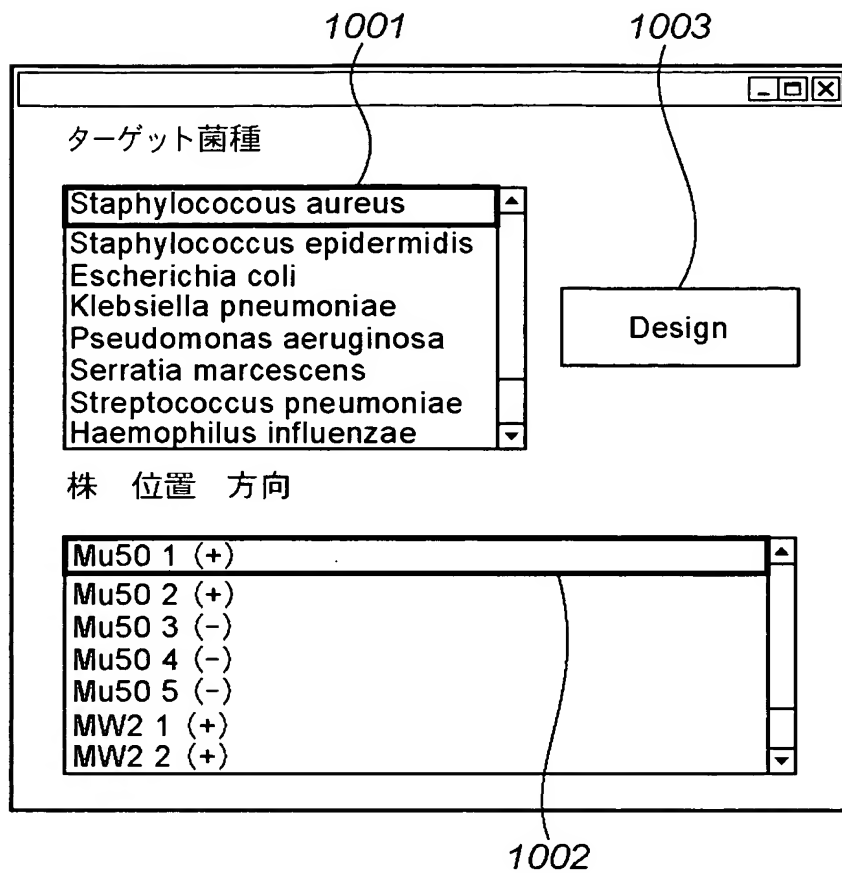
【図 8】



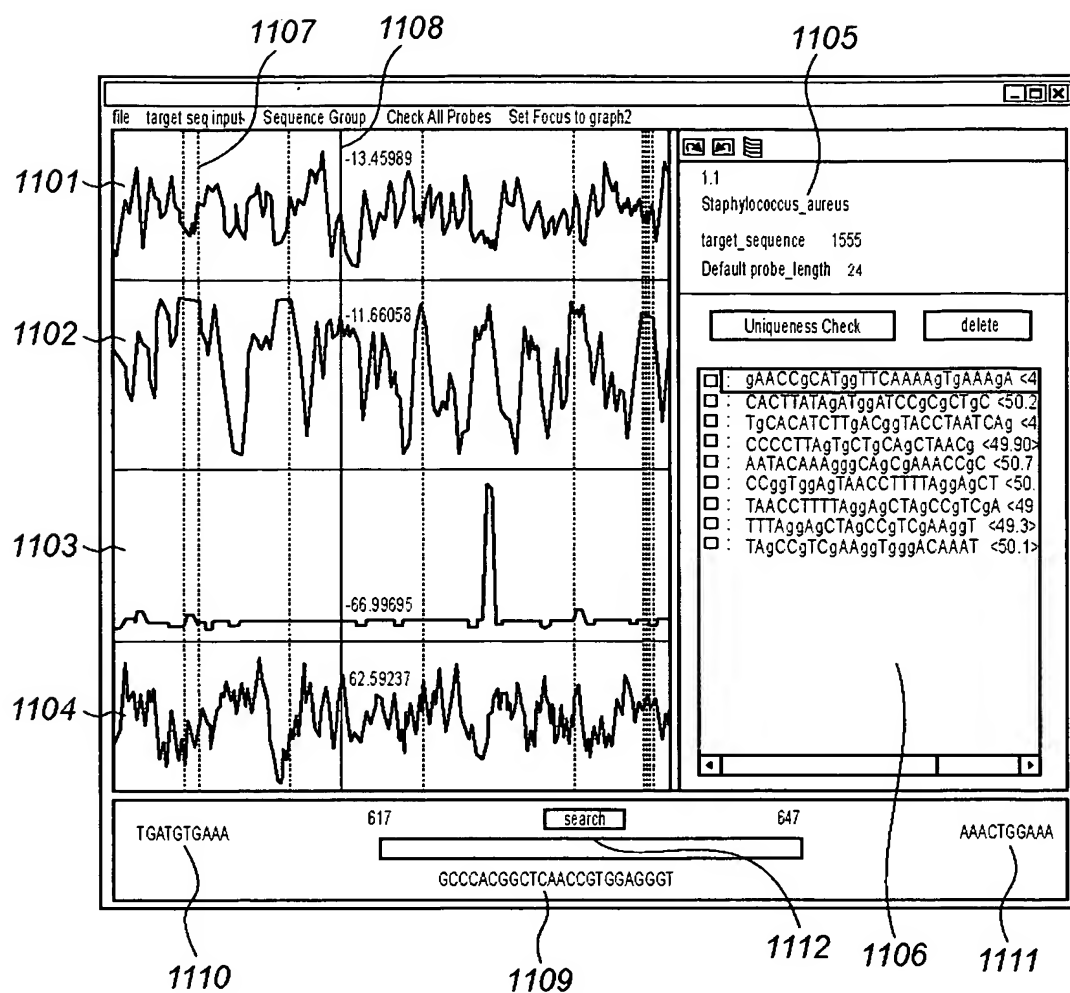
【図 9】



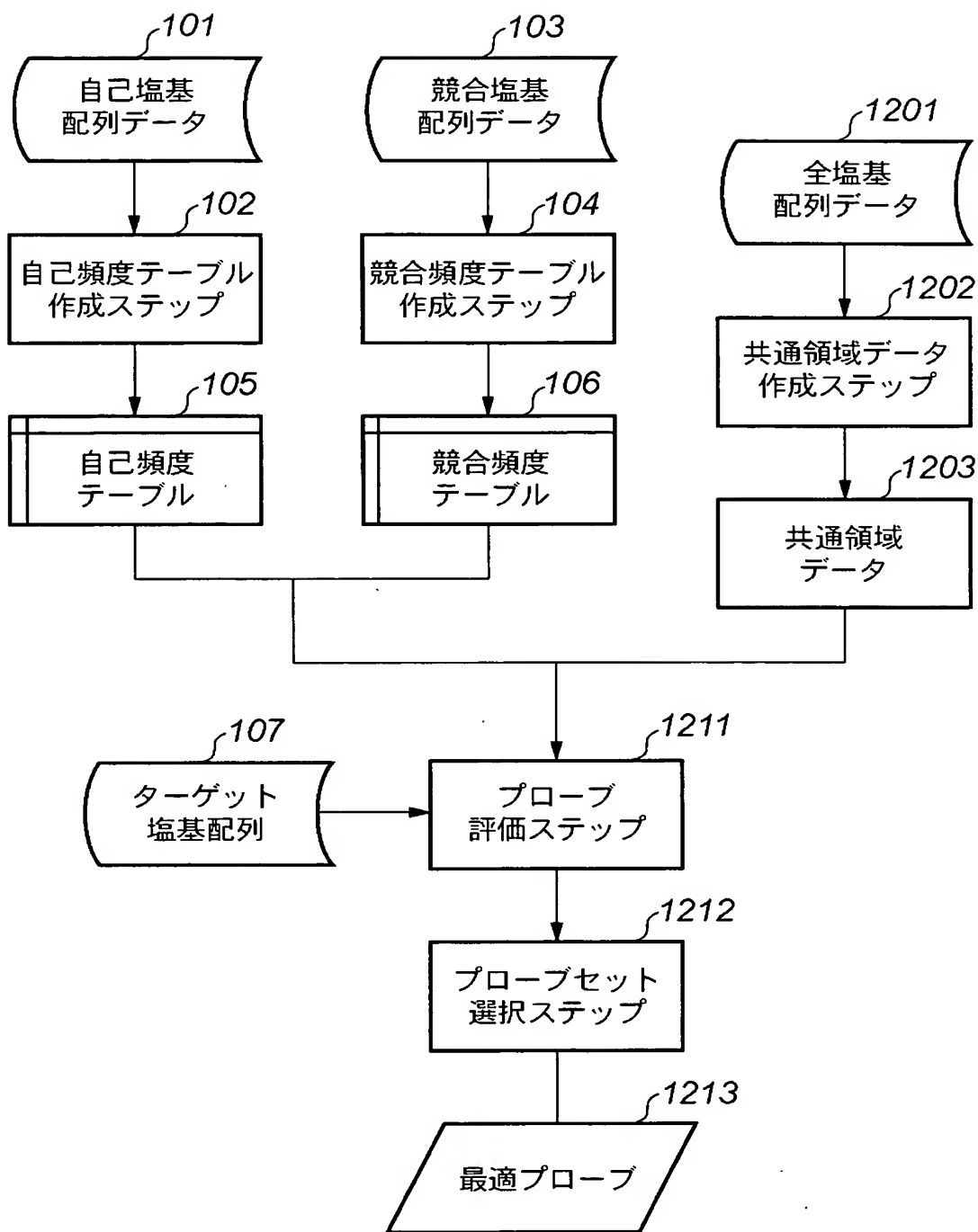
【図 10】



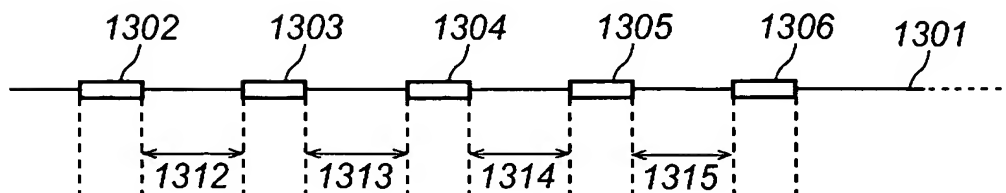
【図 11】



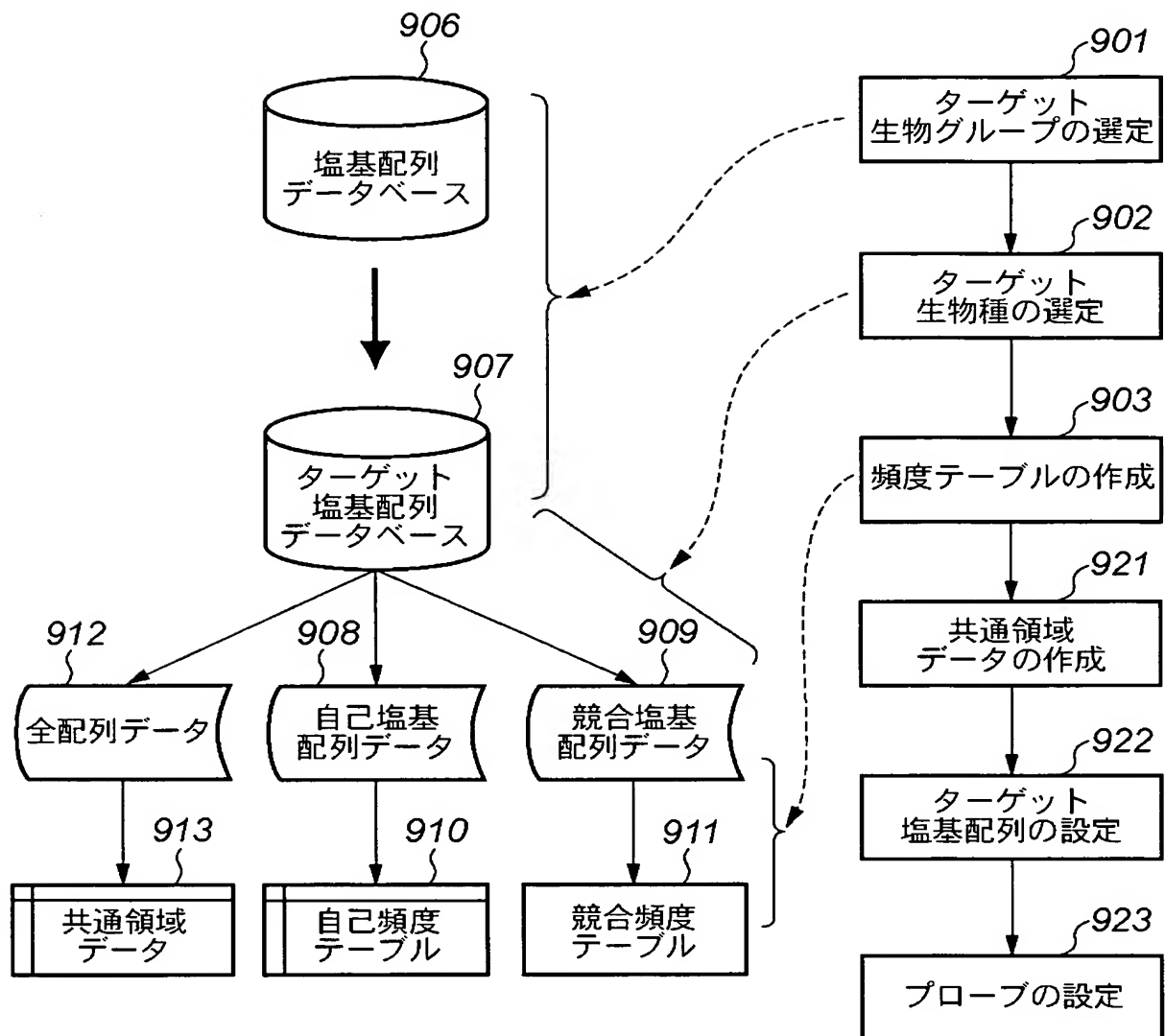
【図 12】



【図 13】



【図 14】





【図 15】

1501

View\_multi\_sequence  
ファイル(F) オプション(O) ヘルプ(H)

学術名	日本語名
▶ Staphylococcus epidermidis	黄色ブドウ球菌
Escherichia coli	大腸菌
Klebsiella pneumoniae	クレブシエラ (肺炎桿菌)
Pseudomonas aeruginosa	緑膿菌
Serratia marcescens	セラチア (膿菌)
Streptococcus pneumoniae	肺炎レンサ球菌
Haemophilus influenzae	インフルエンザ桿菌

.tattggcg	..... lagcggt
tattggcg	..... lagcggt
.tattggcg	..... lagcggt
tattggcg	..... lagcggt
tattggcg	..... lagcggt
tattggcg	..... lagcggt
tattggcg	..... lagcggt
tattggcg	..... lagcggt
tattggcg	..... lagcggt
tattggcg	..... lagcggt
tattggcg	..... lagcggt
tattggcg	..... lagcggt

1502

学術名: Staphylococcus aureus  
株名: MU50  
株ID: 1827  
Number\_16s: 0

ACCESSION: Y15856.1  
GI: 2950318

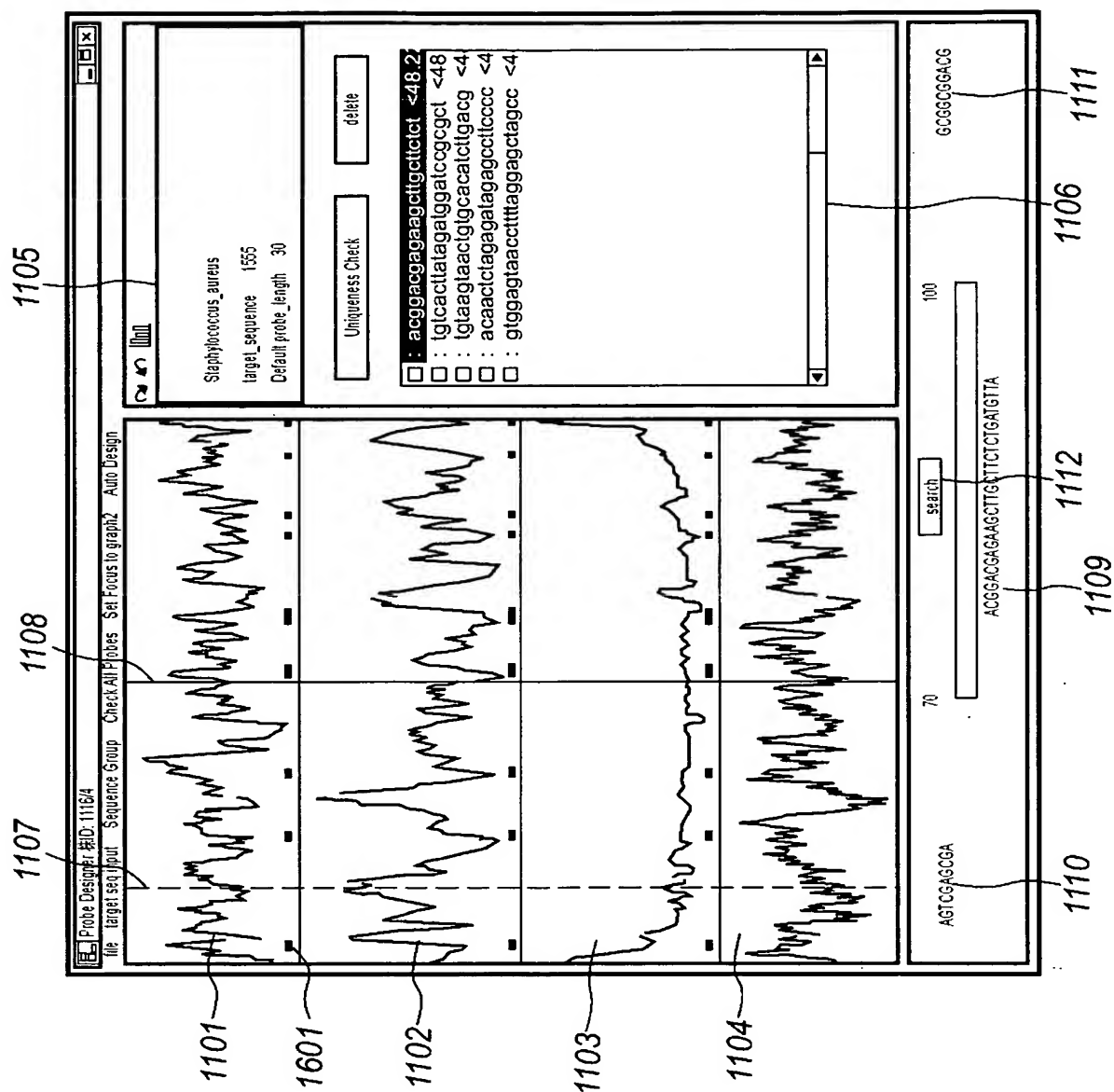
Probe Designer...

1503

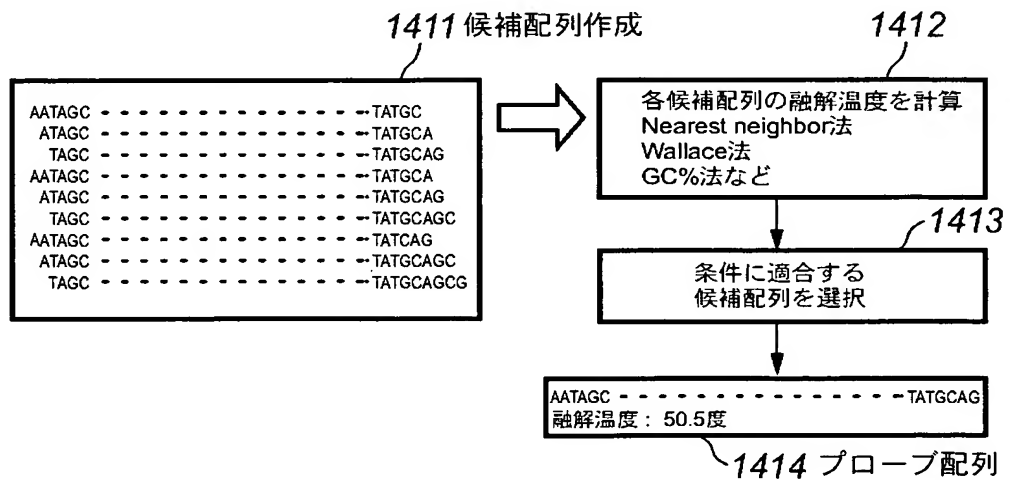
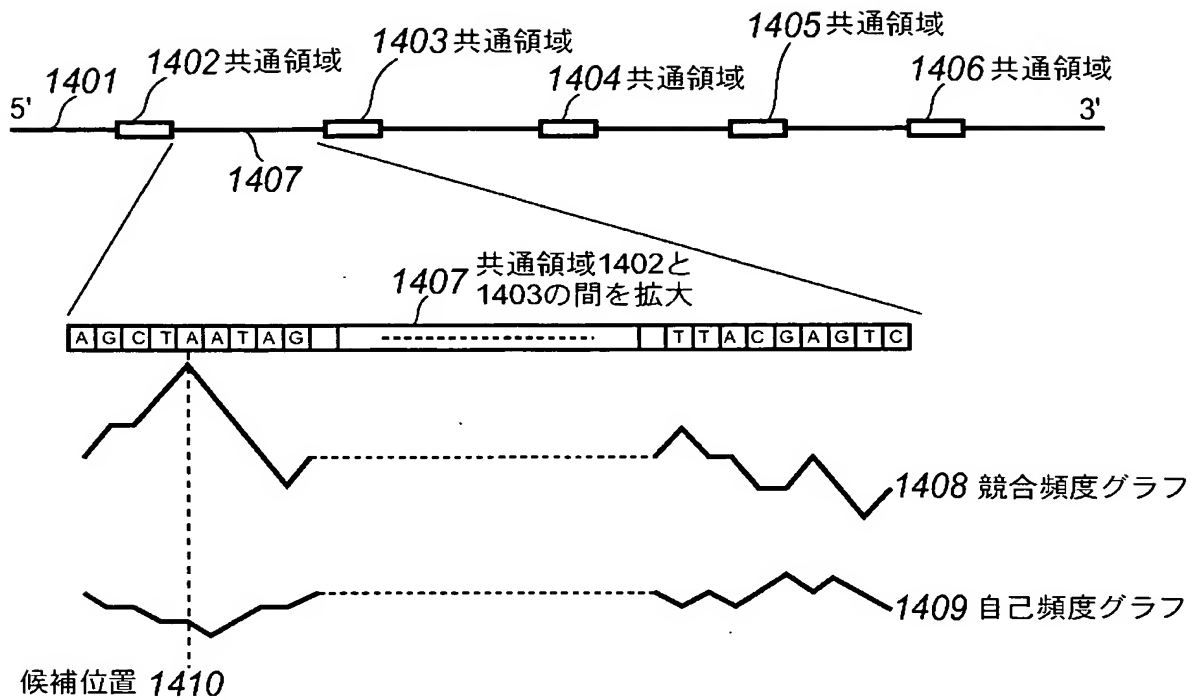
Staphylococcus aureus (黄色ブドウ球菌)

1504

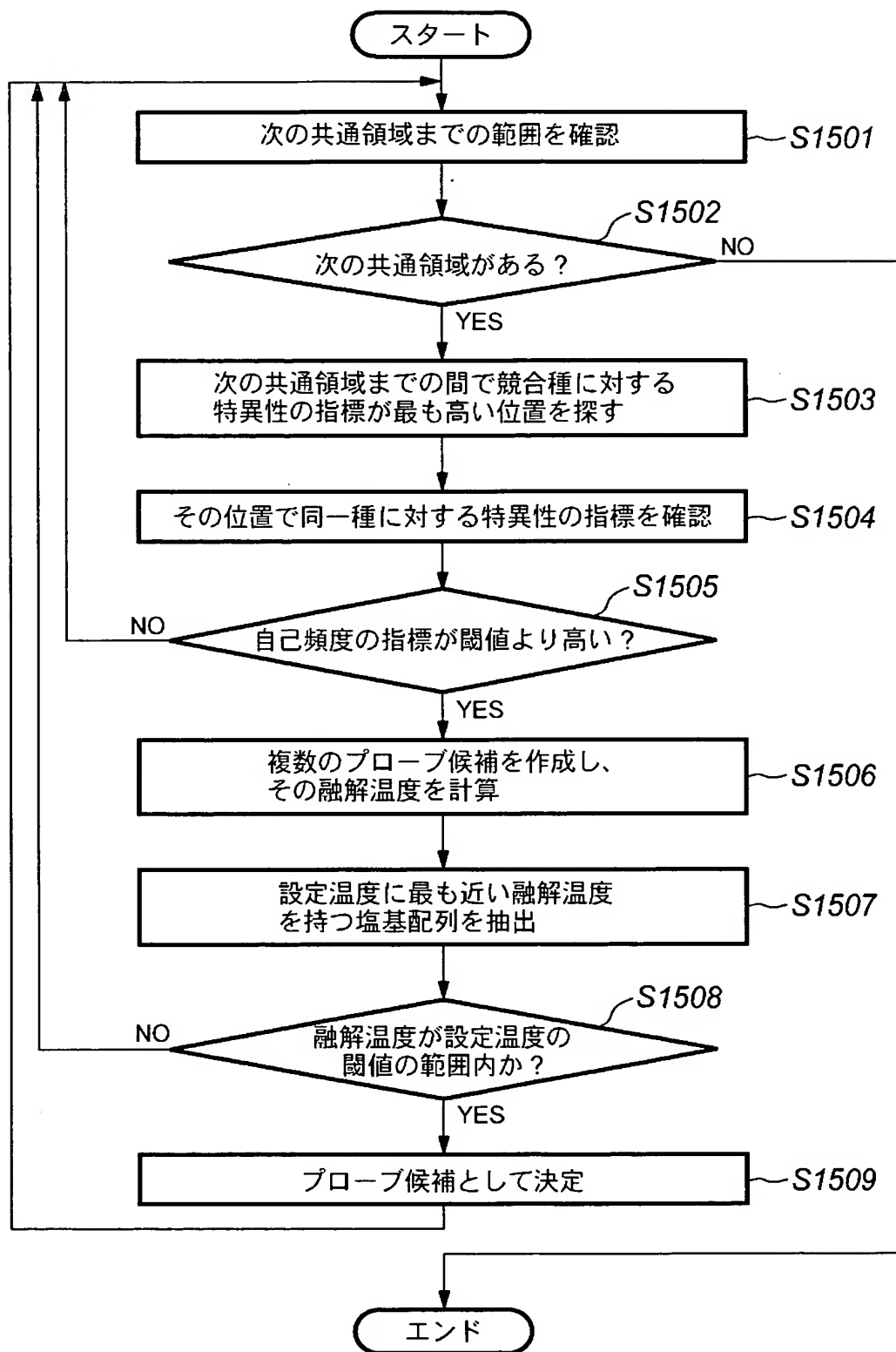
【図 16】



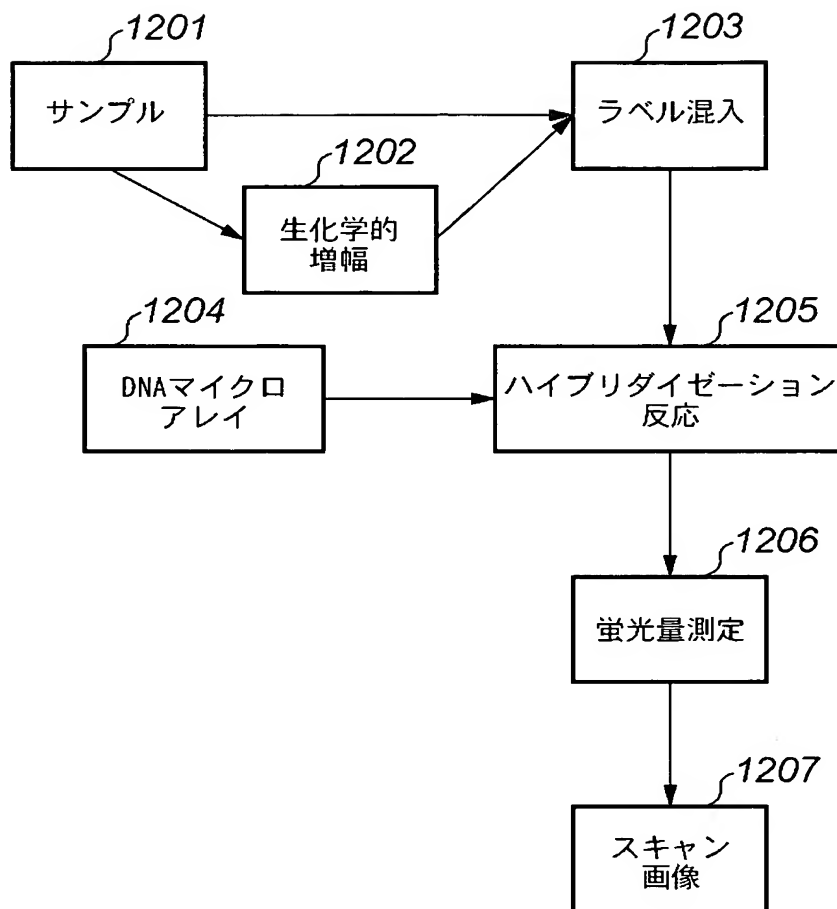
【図 17】



【図 18】



【図 19】



**【書類名】 要約書****【要約】**

**【課題】** 正確かつ、再現性のある DNA プローブ設計を可能とする。

**【解決手段】** ターゲット塩基配列を含む自己塩基配列データ 1 0 1 について、ターゲット塩基配列 1 0 7 から得られる複数の部分塩基配列の各々の出現数をカウントして得られる自己頻度テーブル 1 0 5 を保持する。また、自己塩基配列データ 1 0 1 とは区別されるべき競合塩基配列データ 1 0 3 について、上記複数の部分塩基配列の各々の出現数をカウントして得られる競合頻度テーブルを保持する。プローブ評価ステップ 1 0 8 では、自己及び競合頻度テーブル 1 0 5、1 0 6 の頻度情報を部分塩基配列を基準に対比可能に提示し、ユーザによる指示操作に応じて複数の部分塩基配列の少なくとも一つを特定し、特定された部分塩基配列に基づいてプローブ候補が形成される。

**【選択図】** 図 1

## 認定・付加情報

特許出願の番号	特願 2004-064494
受付番号	50400380368
書類名	特許願
担当官	第七担当上席 0096
作成日	平成 16 年 3 月 11 日

## &lt; 認定情報・付加情報 &gt;

## 【特許出願人】

## 【識別番号】

000001007

## 【住所又は居所】

東京都大田区下丸子 3 丁目 30 番 2 号

## 【氏名又は名称】

キヤノン株式会社

## 【代理人】

申請人

## 【識別番号】

100076428

## 【住所又は居所】

東京都千代田区紀尾井町 3 番 6 号 秀和紀尾井町  
パークビル 7 F 大塚国際特許事務所

## 【氏名又は名称】

大塚 康德

## 【選任した代理人】

## 【識別番号】

100112508

## 【住所又は居所】

東京都千代田区紀尾井町 3 番 6 号 秀和紀尾井町  
パークビル 7 F 大塚国際特許事務所

## 【氏名又は名称】

高柳 司郎

## 【選任した代理人】

## 【識別番号】

100115071

## 【住所又は居所】

東京都千代田区紀尾井町 3 番 6 号 秀和紀尾井町  
パークビル 7 F 大塚国際特許事務所

## 【氏名又は名称】

大塚 康弘

## 【選任した代理人】

## 【識別番号】

100116894

## 【住所又は居所】

東京都千代田区紀尾井町 3 番 6 号 秀和紀尾井町  
パークビル 7 F 大塚国際特許事務所

## 【氏名又は名称】

木村 秀二

特願 2 0 0 4 - 0 6 4 4 9 4

出 願 人 履 歴 情 報

識別番号

[ 0 0 0 0 0 1 0 0 7 ]

1. 変更年月日

1 9 9 0 年 8 月 3 0 日

[変更理由]

新規登録

住 所

東京都大田区下丸子 3 丁目 3 0 番 2 号

氏 名

キャノン株式会社